International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
&y = \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right),\\
&y_{-1} = \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \qquad (1)\\
&x = \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned}
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \tag{1}$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \tag{2}$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2(W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \tag{3}$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \tag{4}$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = (XW\hat{V}W'X)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i'F_T\hat{V}_i\hat{V}_i'F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = (XW\hat{V}W'X)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = {e'e}/{(NT-2)}$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2(W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
& y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\
& y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\
& x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})
\end{aligned} \tag{1}
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$
\gamma = (W'W)^{-1} W'y. \tag{2}
$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = {e'e}/{(NT-2)}$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$
y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \tag{3}
$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$
\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \tag{4}
$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15.
Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15.
Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = {e'e}/{(NT-2)}$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15.
Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1} \mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15.
Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\ y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\ x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right)x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
&y = \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
&y_{-1} = \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), \textit{and} \\
&x = \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \qquad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}, \text{with } s^2 = {e'e}/{(NT-2)}$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \quad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \quad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \quad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \quad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15.
Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1} x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1} \mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned}
\qquad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned}
\tag{1}
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \tag{2}$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \tag{3}$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \tag{4}$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	OLS	ABGMM	RMSE (β) SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15.
Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), \, and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}, \text{with } s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \tag{1}$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \tag{2}$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2(W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \tag{3}$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \tag{4}$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$
\begin{aligned}
y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),\\
y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\
x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})
\end{aligned} \quad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(E(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), \, and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$
\begin{aligned}
& y=\left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
& y_{-1}=\left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
& x=\left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \qquad (1)
$$

Also, let $W=\left[y_{-1}x\right]$. Then the OLS estimator of the parameter vector $(\alpha\beta)'=\gamma$ is given by

$$\gamma=\left(W'W\right)^{-1}W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma)=s^2\left(W'W\right)^{-1}$, with $s^2={e'e}/{(NT-2)}$, where $e=(y-W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $\left(W'W\right)^{-1}W'diag(e'e)W\left(W'W\right)^{-1}$. Since $Cov\left(y_{i,t-1}\mu_i \neq 0\right)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it}-y_{i,t-1}=\delta\left(y_{i,t-1}-y_{i,t-2}\right)\left(x_{it}'+x_{i,t-1}'\right)\beta+v_{it}+v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $\left(\mathrm{E}\left(x_{it}'-\mu\right)\neq 0\right)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2}-y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1}-y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH}=\left(XPX\right)^{-1}X'Py \quad \text{Where} \quad P=Z\left(Z'Z\right)^{-1}Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\ y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\ x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right) \end{aligned} \qquad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2(W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1}Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right),$$
$$y_{-1} = \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \qquad (1)$$
$$x = \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i,t-2}\right)\left(x'_{it} + x'_{i,t-1}\right)\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $\left(\mathrm{E}(x'_{it} - \mu) \neq 0\right)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = (XW\hat{V}W'X)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
&y=\left(y_{11},\Lambda,y_{N1},\Lambda,y_{iT},\Lambda,y_{NT}\right),\\
&y_{-1}=\left(y_{10},\Lambda,y_{N0},\Lambda,y_{i,T-1},\Lambda,y_{N,T-1}\right),and \qquad (1)\\
&x=\left(x_{10},\Lambda,x_{N0},\Lambda,x_{i,T-1},\Lambda,x_{N,T-1}\right)
\end{aligned}
$$

Also, let $W=\left[y_{-1}x\right]$. Then the OLS estimator of the parameter vector $(\alpha\beta)'=\gamma$ is given by

$$\gamma=\left(W'W\right)^{-1}W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma)=s^2\left(W'W\right)^{-1}$, with $s^2=e'e/(NT-2)$, where $e=(y-W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $\left(W'W\right)^{-1}W'diag(e'e)W\left(W'W\right)^{-1}$. Since $Cov(y_{i,t-1}\mu_i\neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it}-y_{i,t-1}=\delta\left(y_{i,t-1}-y_{i,t-2}\right)\left(x'_{it}+x'_{i,t-1}\right)\beta+v_{it}+v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $\left(\mathrm{E}\left(x'_{it}-\mu\right)\neq 0\right)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2}-y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1}-y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH}=\left(XPX\right)^{-1}X'Py \quad \text{Where} \quad P=Z\left(Z'Z\right)^{-1}Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)},\hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$\begin{aligned} y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\ y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\ x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right) \end{aligned} \tag{1}$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \tag{2}$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \tag{3}$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \tag{4}$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \tag{1}
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y . \tag{2}$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i,t-2}\right)\left(x'_{it} + x'_{i,t-1}\right)\beta + v_{it} + v_{i,t-1} \tag{3}$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $\left(\mathrm{E}(x'_{it} - \mu) \neq 0\right)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \tag{4}$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
&y = \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
&y_{-1} = \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
&x = \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \quad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT - 2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = (XW\hat{V}W'X)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i'F_T\hat{V}_i\hat{V}_i'F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \quad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y . \quad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i'F_T\hat{V}_i\hat{V}_i'F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right),\\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \qquad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = {e'e}/{(NT-2)}$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\
y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\
x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})
\end{aligned} \quad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), \textit{and} \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i'F_T\hat{V}_i\hat{V}_i'F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \quad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \quad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \quad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \quad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
&y = \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
&y_{-1} = \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \qquad (1) \\
&x = \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned}
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right),\\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \qquad (1)\\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned}
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), \text{ and} \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right)x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1} \mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), \textit{and} \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}, \text{with}\ s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \qquad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}),\ and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1} \mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i'F_T\hat{V}_i\hat{V}_i'F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \tag{1}
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \tag{2}$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \tag{3}$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \tag{4}$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right)x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned}
\qquad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), \textit{and} \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \qquad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$
\gamma = (W'W)^{-1} W'y. \qquad (2)
$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$
y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i,t-2}\right)\left(x'_{it} + x'_{i,t-1}\right)\beta + v_{it} + v_{i,t-1} \qquad (3)
$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$
\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)
$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2(W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$
\begin{aligned}
y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\
y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\
x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right)
\end{aligned} \quad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), \text{ and} \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right)x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), \; and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15.
Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = (XW\hat{V}W'X)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1) T	δ	ρ	RMSE (β) OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT - 2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}, \text{with } s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

IJR

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

$$
\begin{aligned}
y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),\\
y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and\\
x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})
\end{aligned} \qquad (1)
$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = {e'e}/{(NT-2)}$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = (X'WX)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}), \\ y_{-1} &= (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \\ x &= (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2(W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = {e'e}/{(NT-2)}$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \qquad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \qquad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \qquad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \qquad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i . \qquad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

$$\begin{aligned} y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\ y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\ x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right) \end{aligned} \qquad (1)$$

Also, let $W = \left[y_{-1}x\right]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = \left(W'W\right)^{-1} W'y . \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\operatorname{var}(\gamma) = s^2\left(W'W\right)^{-1}$, with $s^2 = e'e \big/ (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $\left(W'W\right)^{-1} W' diag(e'e) W \left(W'W\right)^{-1}$. Since $Cov\left(y_{i,t-1}\mu_i \neq 0\right)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i,t-2}\right)\left(x'_{it} + x'_{i,t-1}\right)\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $\left(\mathrm{E}\left(x'_{it} - \mu\right) \neq 0\right)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = \left(XPX\right)^{-1} X'Py \quad \text{Where} \quad P = Z\left(Z'Z\right)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1)}: \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it}. In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address: uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \qquad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \qquad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \qquad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \qquad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.
Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.
Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.
Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.
Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \qquad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \qquad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \qquad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \qquad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research

p-ISSN: 2348-6848

e-ISSN: 2348-795X

Volume 08 Issue 02

February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\mathrm{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e / (NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W' diag(e'e) W (W'W)^{-1}$. Since $Cov(y_{i,t-1} \mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta\left(y_{i,t-1} - y_{i.t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta\left(y_{i2} - y_{i1}\right) + \left(x'_{i3} - x'_{i2}\right)\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta\left(y_{i3} - y_{i2}\right) + \left(x'_{i4} - x'i_3\right)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where $\mu_i \sim$ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$y = (y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}),$$
$$y_{-1} = (y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}), and \quad (1)$$
$$x = (x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1})$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\text{E}(x'_{it} - \mu) \neq 0)$ and it results in the "loss" of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)})$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^1{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5, 0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3, 0.5, 0.7)$, $\lambda = (0.3, 0.5, 0.7)$, $\rho = (0.2, 0.5, 0.9)$ and $\theta = (0.2, 0.5, 0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

A Monte-Carlo study of Dynamic Panel Data Estimators with Autocorrelated Error Terms

[1]Uthman, Kafayat Tolani and [2]Oyenuga, Iyabode Favour

[1]National Centre for Genetic Resources and Biotechnology, Moor Plantation, Ibadan, Nigeria.
[2]Department of Mathematics and Statistics, The Polytechnic, Ibadan, Nigeria.

Email Address:uthman_tk@yahoo.com , ioyenuga2007@polyibadan.edu.ng

ABSTRACT

Many economic relationships are dynamic in nature and one of the advantages of panel data is that they allow the researcher to understand the dynamics of adjustment. These dynamic relationships are characterized by the presence of a lagged dependent variable among the regressors. In this research, the properties of some Dynamic Panel Data estimators including Ordinary Least Squares were investigated; the Anderson-Hsiao (AH), Arellano-Bond Generalized Method of Moment (GMM) one-step, Blundell- Bond System (SYS1) one-step, M and MM estimators in the presence of serial correlation. Monte-Carlo simulations were carried out at varying sample sizes with different degrees of autocorrelation. The results showed that in small and large sample situations, irrespective of time dimension, Anderson-Hsiao estimator (AH) outperformed all other estimators. Meanwhile, Blundell-Bond System estimator has the least performance among all the estimators.

Keyword: Dynamic panel data, Monte Carlo Simulation, Autocorrelation and Time series data.

1. INTRODUCTION

Panel data or longitudinal data refer to a cross-section repeatedly sampled over time, but where the same economic agent has been followed throughout the period of the sample. It is thus a pooling of observation on a cross-section of individuals, households, firms, regions or countries on several time periods. Because panel data have both cross-sectional and time series dimensions, the application of regression models to fit econometric models are complex than

those for simple cross-sectional or time-series data sets. Nevertheless, they are increasingly being used in applied work.

A number of studies on dynamic panel data modeling have been carried out and reported in the literature based on Instrumental Variable and Generalized Method of Moments (GMM) estimators.(Nickell,1981; Anderson & Hsiao, 1981,1982; Kiviet,1995; Arellano & Bond, 1991; Ahn & Schmidt,1995; Islam,1998;). Dynamic models are of interest in a wide range of economic applications such as Euler equations for household consumption, democracy and education, empirical model of economic growth and so on (Bond, 2002). Estimation of dynamic panel models is unfortunately problematic. For the F Fixed Effects specification, the problem arises as a consequence of relatively short time series component, typical of most panel data sets. Thus, the usual Hurwicz type bias is instigated into Ordinary Least Squares (OLS) estimation of a Fixed Effects dynamic panel model (Nickel, 1981). In the random effects specification, traditional (feasible) Generalized Least Squares (GLS) estimators are similarly biased due to a correlation between the equations disturbance terms (via the individual effect) and lagged variable (Sevestre and Trognon, 1985).

The main purpose of this paper is to introduce serial correlation in a random effects one-way error component model and compare performance of the some of the existing mainstream Instrumental Variable (IV)/ Generalized Method of Moments (GMM) estimators. Various studies had been carried out using Monte-Carlo simulations and small samples, (for example, Judson & Owen, 1999; Hayakawa, 2008; Islam, 1998; Harris and Matyas, 2010 ;). Therefore, this study will examine the performance of different estimators from small samples to large samples with different time dimension.

2. MATERIALS AND METHODS

This work considers one-way error component model with presence of serial correlation in a random effects. The different degree of autocorrelation was introduced via random effects one-way error component model and the coefficient of the serial correlation is taken to be mild, moderate and high. This is in line with the works of Islam (1998), Judson and Owen (1999), Harris and Matyas (2010) to mention but few. Most of the previous works done on Dynamic panel data were focused on the absence or no serial correlation of the disturbance term.

2.1 BRIEF OVERVIEW OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS CONSIDERED

Ordinary Least Square estimators are applied to the equation in level form. Let

$$\begin{aligned} y &= \left(y_{11}, \Lambda, y_{N1}, \Lambda, y_{iT}, \Lambda, y_{NT}\right), \\ y_{-1} &= \left(y_{10}, \Lambda, y_{N0}, \Lambda, y_{i,T-1}, \Lambda, y_{N,T-1}\right), and \\ x &= \left(x_{10}, \Lambda, x_{N0}, \Lambda, x_{i,T-1}, \Lambda, x_{N,T-1}\right) \end{aligned} \quad (1)$$

Also, let $W = [y_{-1}x]$. Then the OLS estimator of the parameter vector $(\alpha\beta)' = \gamma$ is given by

$$\gamma = (W'W)^{-1} W'y. \quad (2)$$

The standard errors under homoscedasticity are obtained from $\text{var}(\gamma) = s^2 (W'W)^{-1}$, with $s^2 = e'e/(NT-2)$, where $e = (y - W\gamma)$. The general heteroskedasticity consistent standard errors are obtained from $(W'W)^{-1} W'diag(e'e)W(W'W)^{-1}$. Since $Cov(y_{i,t-1}\mu_i \neq 0)$ OLS estimator is biased. It is also inconsistent in direction of both N and T.

Anderson and Hsiao (1981) proposed an instrumental Variable (IV) estimator that is consistent for fixed T and N tends to infinity. Anderson and Hsiao (IV) estimate apply to the model in first differenced form

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i,t-2})(x'_{it} + x'_{i,t-1})\beta + v_{it} + v_{i,t-1} \quad (3)$$

Which cancelled the individual fixed effects assumed to possibly correlate with the exogenous variables $(\mathrm{E}(x'_{it} - \mu) \neq 0)$ and it results in the “loss” of one cross-section from the actual estimation.

They also suggest use of level Instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$.

The Anderson-Hsiao estimator is given as

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{Where} \quad P = Z(Z'Z)^{-1} Z \quad (4)$$

The symbol l or d to indicate the use of levels or differences as instrument $\left(\hat{\gamma}^{AH(l)}, \hat{\gamma}^{AH(d)}\right)$

Arellano and Bond estimator is similar to the one suggested by Anderson-Hsiao but exploits additional moment restrictions, which enlarges the set of instruments. The dynamic equation to

be estimated in levels is $y_{it} = \delta y_{i,t-1} + x'_{it}\beta + \mu_i + v_{it}$ where the individual effect μ_i is eliminated by differencing

$$y_{it} - y_{i,t-1} = \delta(y_{i,t-1} - y_{i.t-2}) + (x'_{it} - x'_{i,t-1})\beta + v_{it} - v_{i,t-1}. \quad (5)$$

We look for the instruments available for instrumenting the difference equation for each year. For t=3 the equation to be estimated is

$$y_{i3} - y_{i2} = \delta(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + v_{i3} - v_{i2} \quad (6)$$

Where the instruments (again assuming x being at least predetermined) y_{i1}, x'_{i2} and x'_{i1} are available. For t=4 the equation is

$$y_{i4} - y_{i5} = \delta(y_{i3} - y_{i2}) + (x'_{i4} - x'i_3)\beta + v_{i4} - vi_3. \quad (7)$$

The instruments $y_{i,1}, yi_{,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available.

The Arellano-Bond estimator is given by

$$\hat{\gamma}^{ABGMM} = \left(XW\hat{V}W'X\right)^{-1} X'W\hat{V}^{-1}W'Y. \quad (8)$$

Where the one-step GMM estimator makes use of a covariance matrix taking autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W'G_T W_i \quad (9)$$

The two-step GMM estimator makes use of the residuals of the first-step estimation to estimate the covariance matrix as suggested by white (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{V}_i \hat{V}_i' F_T W_i. \quad (10)$$

The Blundell and Bond System GMM: When the instruments are weak the GMM estimator suggested by AB is known to be rather inefficient because making use of the information contained in differences only. The BB suggests making use of additional level information beside the differences to make it an efficient estimator.

The Blundell-Bond estimator is given by

$$\hat{\gamma}^{GMM-SSY} = \left(XW\hat{V}^{-1}W'X\right)^{-1} X'W\hat{V}^{-1}Wy. \qquad (11)$$

The first step GMM estimator makes use of a covariance matrix taking this autocorrelation into account, enlarged for the level equations while the second step GMM estimator uses residuals of the first step estimation to estimate the matrix as suggested by white (1980).

M-estimators were proposed by [13]. M-estimation for regression is a relatively straightforward extension of M- estimation for location and scale. It represents one of the first attempts at a compromise between the efficiency of the Least Squares estimator and a resistance estimator - Least Absolute Value (LAV) estimators. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to the M-estimates nonlinear normal equations.

$$\sum_{i=1}^{n} \Psi\left(\frac{y_i - x'\beta_1}{s}\right) x_i = 0 \qquad (12)$$

IRLS express the normal equations as

$$X'WX\hat{\beta} = X'Wy \text{ where W is an n x n diagonal matrix of weights} \qquad (13)$$

The initial vector of parameter estimates $\hat{\beta}_0$ are typically obtained from OLS,

M estimator is given as

$$\hat{\beta}_1 = \left(X'WX\right)^{-1} X'Wy \qquad (14)$$

MM estimation was introduced by [21] and it combines a high breakdown point with a good efficiency (approximately 95%) relative to the Ordinary Least Squares estimator under the Gauss-Markov assumptions. The MM refers to the fact that Multiple M-estimation procedure is used to calculate the final estimate. It has also become most common robust regression technique for linear regression.

2.2 SIMULATION STUDY

Monte-Carlo experiments are carried out to compare the behaviour of different estimators under different circumstances. The data generating process follows Nerlove (1971) and Arellano- Bond (1991).

$$y_{it} = \delta y_{i,t-1} + X^{1}{}_{it}\beta + \mu_i + v_{it} \quad (15)$$

$$x_{it} = \lambda x_{i,t-1} + \varepsilon_t \quad (16)$$

Where $xit \approx u(-0.5,0.5)$

For the random effects specification, we generate $u_{it} = \mu_i + v_{it}$ where μ_i ~ N (0, 1) and error term v_{it} is generated by

$$\text{AR (1):} \quad v_{it} = \rho v_{i,t-1} + w_{it} \quad (17)$$

Or by the MA (1) process

$$v_{it} = w_{it} + \theta w_{i,t-1} \quad (18)$$

The parameter δ and β is given as $\gamma = (\delta, \beta)^1$ for the value of $\beta = 1$.the parameters that are varied in the simulation are autoregressive coefficient δ and λ and the autocorrelation coefficient ρ and θ.The values of $\delta = (0.3,0.5,0.7)$, $\lambda = (0.3,0.5,0.7)$, $\rho = (0.2,0.5,0.9)$ and $\theta = (0.2,0.5,0.9)$. We choose cross sectional units, N= (10, 20, 50) and Time periods, T= (5, 10, 15, 20).For each combination of N and T with 1000 replications. The assessments of the various estimators considered in this work were based on the absolute bias and RMSE of parameter estimates.

3. RESULTS AND DISCUSSIONS

These results of the performance of the estimators were considered at various levels of autocorrelation. These estimators were ranked using the ranks 1, 2,3,4,5, and 6 with rank 1 to the best estimator that has lowest value of the absolute bias and root mean square error. A rank 2 is assigned to the second best estimator and a rank 3 is assigned to the third best estimator and so on. Table 1 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for δ, Table 2 show the sum of ranks of the estimators using RMSE criterion when N=200 at different time period for δ, Table 3 show the sum of ranks of the estimators using Bias criterion when N=10 at different time period for δ, Table 4 show the sum of ranks of the estimators using Bias criterion when N=200 at different time period for δ, while Table 5 show the sum of ranks of the estimators using RMSE criterion when N=10 at different time period for β.

3.1 DISCUSSION

Our simulation result for the estimates of δ (in terms of RMSE), when N=10, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 while Arellano- Bond GMM estimator performs better in some cases when T=5. When N=20, Anderson-Hsiao (IV) estimator outperforms all other estimator at T=5 and Blundell-Bond System GMM performs better at T=10.But as when N increases, Anderson-Hsiao (IV) estimator outperforms all other estimators at T=5 and 10 respectively.

For estimate of β, we find out that when N is small, Arellano-Bond GMM estimator outperforms all other estimators at T=5 and 10 expect in some cases while M and MM estimators also performs well. As N increases to 50, Arellano-Bond GMM, M, MM and Anderson-Hsiao (IV) estimator performs quite well. However, When N=100 and 200, Ordinary Least Square estimator outperforms all other estimators at T=5, but as T increases to 10, Arellano –Bond GMM estimator performs better than all other estimators.

In terms of bias, Arellano-Bond GMM estimator outperforms all other estimators followed by MM, M, Ordinary Least Square and Anderson-Hsiao (IV) estimators while Blundell-Bond System GMM has worst performance among the estimators.

When the value of the autoregressive parameter of the explanatory variable λ, is varies at (0.3, 0.5 and 0.7). For the parameter δ the bias and RMSE of the estimators improves as the value of λ increases expect Ordinary Least Square and the robust estimators that deteriorates with increase in the value at various combination of N and T. But for the parameter β, the bias and RMSE of all the estimators improves as λ increases.

.

4. CONCLUSIONS

For the estimate of δ, we noted that the estimator proposed by Anderson-Hsiao outperforms all other estimators under all different generating mechanism of v_{it} . In term of bias, Arellano-Bond GMM estimator proves to be relatively superior among the estimators when T is small, though it may not be the best but its performance improved drastically as N increases. The results showed that in small and large sample situations, irrespective of time dimension, Anderson–Hsiao estimator (AH) outperforms all other estimators.

For the estimate of δ, the GMM system estimator proposed by Blundell-Bond System estimator has the least performance when T is small, but as T increases, the estimator improves in the terms of both RMSE and bias.

For most of the estimators, the performance does not vary to great extent with respect to the alternative generating scheme of v_{it}

REFERENCES

Ahn, S.C and P.Schmidt (1995). Efficient Estimation of Models for Dynamic Panel Data. Journal of Econometrica, 68.5-27.

Anderson, T.W and Hsiao, C. (1981). Estimation of Dynamic models with error components. Journal of the American Statistical Association, 76, 598-606.

Anderson, T. and Hsiao, C. (1982). Formulation and Estimation of Dynamic Models using Panel Data. Journal of Econometrics, 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and Application to Employment Equation. Review of Economics Studies, 58, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error- Components Models. Journal of Econometrics, 68, 29-51.

Baltagi, B.H. (2008). Forecasting with Panel Data. Journal of Forecasting, 27,153-173.

Baltagi, B.H. (2005). Econometrics analysis of Panel data. 3rd edition, John Wiley and Sons Ltd, England.

Blundell, R and Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. Journal of Econometrics, 87, 115-143.

Hayakawa, K. (2008). A Simple Efficient Instrumental Variable Estimator for Panel AR (p) Model when Both N and T are large. Department of Economics, Hitotsubashi University JSPS Research fellow.

Huber,P.J. (1973). Robust regression: Asymptotics, Conjectures and Monte-Carlo. Annals of Statistics, 1, 799-821.

Islam, N. (1998). Small Sample Performance of Dynamic Panel Data Estimators: A Monte Carlo Study on the Basis of Growth Data. Department of Economics, Emory University.

Judson, R. and Owen, L. (1996). Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists. Federal Reserve Board of Governors'.

Judson, R. and Owen, L. (1999). Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. Economics Letters, 65, 9-15. Retrieved http:/www.scienpub.com/reference/33922.

Kiviet, J. (1995). On Bias Inconsistency, Efficiency of Various Estimators in Dynamic Panel

International Journal of Research
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 08 Issue 02
February 2021

Data Models. Journal of Econometrics, 68, 53-78.

Mark N. Harris and Laszlo Matyas (2010). A Comparative Analysis of Different Estimators for Dynamic Panel Data Models. The Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Australia.

Nerlove, M. (1971). Further Evidence in the Estimation of Dynamic Economics Relations from a Time Series of Cross Sections. Economic Studies Quarterly, 39, 383-396.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. Econometrica, 49, 1417-1426.

Yohai, V.J. (1987). High breakdown-point and high efficiency robust estimates of regression Annals of Statistics, 15, 642-656.

Appendix

TABLE 1 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for δ

AR(1)			RMSE (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	17	7	13	3	13	10
		0.5	17	5	10	8	13	10
		0.9	18	9	6	3	15	12
	0.5	0.2	17	6	14	3	13	10
		0.5	17	5	8	8	14	11
		0.9	18	9	6	3	15	12
	0.7	0.2	17	6	12	3	14	11
		0.5	17	5	6	12	13	10
		0.9	18	9	5	4	15	12
10	0.3	0.2	16	6	17	3	11	10
		0.5	17	6	14	3	13	10
		0.9	18	7	10	3	14	11
	0.5	0.2	16	6	17	3	12	9
		0.5	17	6	14	3	13	10
		0.9	18	7	9	3	15	11
	0.7	0.2	17	6	15	3	13	9
		0.5	18	6	12	3	14	10
		0.9	18	6	9	3	15	12

TABLE 2 :Sum of Ranks of Estimators using RMSE Criterion when N=200 at different time period for δ

AR(1)			RMSE(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	10	15	3	11	8
		0.5	18	9	10	3	13	10
		0.9	18	9	6	3	12	15
	0.5	0.2	16	9	15	3	12	8
		0.5	18	7	10	3	14	11
		0.9	18	9	6	3	13	14
	0.7	0.2	17	7	14	3	13	9
		0.5	18	7	9	3	15	11
		0.9	18	9	6	3	12	15
10	0.3	0.2	18	9	6	3	13	14
		0.5	18	9	6	3	14	13
		0.9	18	9	6	3	12	15
	0.5	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	13	14
		0.9	18	9	6	3	13	14
	0.7	0.2	18	9	6	3	12	15
		0.5	18	9	6	3	12	15
		0.9	18	9	6	3	12	15

TABLE 3 :Sum of Ranks of Estimators using Bias Criterion when N=10 at different time period for δ

AR(1)			BIAS (δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	8	3	17	9	10
		0.5	13	9	4	9	12	16
		0.9	9	6	3	16	13	16
	0.5	0.2	16	7	3	16	10	11
		0.5	14	8	4	13	9	15
		0.9	9	4	7	18	15	10
	0.7	0.2	17	8	3	15	8	12
		0.5	13	9	3	15	7	16
		0.9	9	6	3	14	14	17
10	0.3	0.2	12	3	7	18	15	8
		0.5	12	4	8	18	15	6
		0.9	15	4	12	14	6	12
	0.5	0.2	11	4	5	18	15	10
		0.5	11	4	7	18	15	8
		0.9	16	3	11	14	9	10
	0.7	0.2	10	4	5	18	15	11
		0.5	12	4	7	18	15	7
		0.9	15	4	11	15	11	7

TABLE 4 :Sum of Ranks of Estimators using Bias Criterion when N=200 at different time period for δ

AR(1)			BIAS(δ)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	16	4	16	7	13	7
		0.5	18	4	12	5	15	9
		0.9	18	5	4	10	11	15
	0.5	0.2	17	4	12	7	16	7
		0.5	18	5	12	5	15	8
		0.9	18	6	3	9	12	15
	0.7	0.2	17	3	13	7	15	8
		0.5	18	5	12	6	15	7
		0.9	18	6	3	9	12	15
10	0.3	0.2	13	3	12	6	13	16
		0.5	10	3	10	7	18	15
		0.9	10	3	10	7	17	16
	0.5	0.2	13	4	12	7	14	13
		0.5	11	3	10	6	18	15
		0.9	11	3	10	6	18	15
	0.7	0.2	13	3	11	6	13	17
		0.5	11	3	11	6	16	16
		0.9	11	3	10	6	18	15

TABLE 5 :Sum of Ranks of Estimators using RMSE Criterion when N=10 at different time period for β

AR(1)			RMSE (β)					
T	δ	ρ	OLS	ABGMM	SYS 1	AH(d)	M-Est.	MM-Est.
5	0.3	0.2	12	7	15	18	5	6
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.5	0.2	11	8	15	18	7	4
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
	0.7	0.2	11	9	15	18	7	3
		0.5	12	4	15	18	9	5
		0.9	12	3	15	18	9	6
10	0.3	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.5	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12
	0.7	0.2	6	3	15	18	9	12
		0.5	6	3	15	18	9	12
		0.9	6	3	15	18	9	12

Note that the boldfaced number are the one with lowest sum of ranks, are the best estimators.