Fitting of Engel's curve for rural Uttar Pradesh

Sana Samreen

Senior research fellow, Dept. of economics, Aligarh Muslim University, Aligarh.
alif.sofi@gmail.com

Abstract:

This study aims at understanding the trends of food consumption in households with changing income levels. The input data that is used for this purpose has been extracted from the National Sample Survey on Household Consumer Expenditure which was conducted from 1st July 2006 to 30th June 2007. The region that has been chosen for the study is rural Uttar Pradesh. The objective is achieved by plotting an Engel Curve and subsequently analyzing the pattern observed. Ideally, the Engel curve is plotted between Income of the Consumer and the Quantity Consumed. But as Income is directly proportional to Expenditure, we have taken total expenditure as a proxy to income for our analysis. Univariate and multivariate models of analysis have been used in our study. The paper has been divided into four broad parts that take into account the background of the problem in the introductory part, datasource and methodology, methodology and major findings or concluding observations.

Key Words:

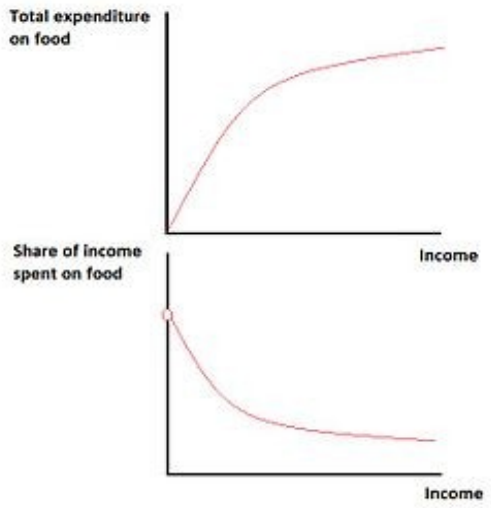
Consumption; expenditure; income levels

I. Introduction:

Food expenditure pattern is an excellent indicator of economic wellbeing of people. If the society is wealthy proportionately high expenditure will be made on secondary necessities, comfort, luxury products and conspicuous consumption. On the other hand, if the society is at subsistence level, people will spend proportionately more on food. Engel's lawthus states that the proportion of income spent on food falls, even if actual expenditure on food rises. In other words, the income elasticity of demand of food is between 0 and 1. The law doesn't imply that food spending remains unchanged as income increases: It suggests that consumers increase their expenditures for food products (in % terms) less than their increases in income.

"One application of this statistic is treating it as a reflection of the living standard of a country. As this proportion or "Engel coefficient" increases, the country is by nature poorer, conversely a low Engel coefficient indicates a higher standard of living."

In the present study, we also tried to acknowledge one very important aspect of elasticity as determined with the help of Engel curves which can throw quite a significant amount of light on the problem of food inflation in India.As in the previous studies too it has been revealed that it is the growing purchasing power of the rural class and consequently their increasing expenditure on food itemswhich is the main driver of increasing food inflation in India.However this is just an observation which can be a crucial aspect of Engel curves.



Multiple regression analysis is used in our study considering various factors as input to arrive at Engel curves. Graphs are plotted for the expenditure on food across two sample populations considering all inputs in different functional forms Linear, Semi Log and Double Log, working-leser and rank 3 leser formsThus, by fitting of the engel curves, not only have we been able to get a picture of the expenditure patterns, but conclusions can also be drawn regarding the inflation levels,elasticity of demand of households ,determination of the standard of living etc.

Engel Curve:

An Engel curve describes how a consumer's purchases of a good like food varies as the consumer's total resources such as income or total expenditures vary. Engel curves may also depend on demographic variables and other consumer characteristics. A good's Engel curve determines its income elasticity, and hence whether the good is an inferior, normal, or luxury good. Empirical Engel curves are close to linear for some goods, and highly nonlinear for others. Engel curves are used for equivalence scale calculations and related welfare comparisons, and determine properties of demand systems such as aggregability and rank.

Using data from Belgian surveys of working class families, Ernst Engel (1857,1895) studied how households expenditures on food vary with income. He found that food expenditures are an increasing function of income and of family size, but that food budget shares decrease with income. This relationship of food consumption to income, known as Engel's law, has since been found to hold in most economies and time periods, often with the function hi for food i close to linear in log(y), where y is the income, wealth, or total expenditures on goods and services.

Engel curve provides a wealth of information on the community's consumption behavior at various levels of total expenditure and for different family compositions. Dandekar and Rath(1971), Rao (1981), and Deaton and Tarrozi (2000) came very close to taking a full view of the Engel curve but confined to look only at the poverty line portion of the Engel curve. In determining the total expenditure at which the minimum calorie requirements are met Dandekarand Rath used the observed empirical relation between food expenditure and the total expenditure. Dandekar and Rath looked at the Engel curve based on the National Sample Survey data to determine the total expenditure that would permit an expenditure on food which would meet the calorie requirements. They used the Engel curve only for that purpose. BhanojiRaoused the typical properties of an Engel curve of a necessity to suggest that the proportion of food expenditure increases, reaches a maximum and then declines. He suggested that the point where this proportion reaches a maximum could be taken as the threshold for acute poverty and suggested that one and half times that level can be taken as the poverty line. Again he left Engel curve behind after deriving a poverty line from it.

The Engel curve for a commodity is actually the demand function for that commodity, keeping the prices constant. As such it depicts the consumption behavior of persons or households with different levels of total expenditure. Deaton, who has done considerable work on consumer behavior and consumer demandfunctions, also examined poverty issue only from the traditional approach and ignored looking at the entire Engel curve for essential commodities. The Engel curve not only suggests how to fix a poverty threshold it also provides a distribution of mean consumption expenditure on food according to income. Thus the Engel curve summarizes the economic equilibrium consumption expenditure on food.

II. Nature and source of data :

The National Sample Survey (NSS), set up by the Government of India in 1950 was aimed at determining socio-economic factors by employing various scientific sampling methods. We have taken the 63rd round data of NSSO. The National Sample Survey Organization conducts nationwide sample surveys on various socio-economic issues. The results of these surveys are released in the form of various reports, which are mainly based on tabulation of data according to a preplanned scheme. A well designed schedule of enquiry can be canvassed out of this survey to ascertain the consumption expenditure of households. . The data has been grouped in to blocks.

As the data was in crude form, so we,after bringing it into a readable form, classified it into various heads as directed in the layout file. The National Sample Survey Organization conducts nationwide sample surveys on various socio-economic issues. The results of these surveys are released in the form of various reports, which are mainly based on tabulation of data according to a preplanned scheme, study of sub samples further help in the analysis.

The National Sample Survey

The National Sample Survey (NSS), set up by the Government of India in 1950 to collect socio-economic data employing scientific sampling methods, started its sixty-third round from 1st July 2006.

The sixty-third round of NSS is earmarked for collection of data on economic and operational characteristics of enterprises in service sector (excluding trade) and also on household consumer expenditure. The field operations of the survey commenced on 1st July 2006 and has continued up to 30th June 2007

Geographical coverage: The survey covers the whole of the Indian Union except (i) Leh (Ladakh) and Kargil districts of Jammu & Kashmir, (ii) interior villages of Nagaland situated beyond five kilometres of the bus route and (iii) villages in Andaman and Nicobar Islands which remain inaccessible throughout the year.

Period of survey and work program: The period of survey was of one year duration starting on 1st July 2006 and ending on 30th June 2007. The survey period of this round is divided into four sub-rounds of three months duration each as follows:

Sub-round 1: July - September 2006
Sub-round 2: October - December 2006
Sub-round 3: January - March 2007
Sub-round 4: April - June 2007
In each of these four sub-rounds, equal number of sample villages/ blocks (FSUs) are allotted for survey with a view to ensuring uniform spread of sample FSUs over the entire survey period. Attempt should be made to survey each of the FSUs during the sub-round to which it has been allotted. Because of the arduous field conditions, this restriction need not be

strictly enforced in Andaman and Nicobar Islands, Lakshadweep, rural areas of Arunachal Pradesh and Nagaland

Sample Design:

Outline of sample design: Two frames have been used for the 63rd round survey viz. List frame and Area frame.

List frame: A list of 1000 service sector companies distributed all over India has been used as list frame. The list of financial sector enterprises has been supplied by RBI. For the other service sector enterprises the list has been supplied by the Ministry of Company Affairs.For all the companies in the list frame, information will be collected considering all the branch offices. A combined schedule 2.345 is to be filled up for the list frame companies covering all the branches. All these companies in the list frame will be surveyed. However, these companies and their branch offices will be excluded from the coverage of the area frame survey to avoid duplication.

There is no sub-round restriction for the list frame unit.

Stratification: Within each district of a State/ UT, two basic strata have been formed: i) rural stratum comprising of all rural areas of the district and (ii) urban stratum comprising of all the urban areas of the district. However, if there are one or more towns with population 10 lakhs or more as per population census 2001 in a district, each of them will also form a separate basic stratum and the remaining urban areas of the district will be considered as another basic stratum. There are 27 towns with population 10 lakhs or more at all-India level as per census 2001.

Sub-stratification for area frame:

Rural sector: If 'r' be the sample size allocated for a rural stratum, the number of substrata formed was 'r/2'. The villages within a district as per frame was first arranged in ascending order of population. Then sub-strata 1 to 'r/2' have been demarcated in such a way that each sub-stratum comprised a group of villages of the arranged frame and have more or less equal population.

III. Methodology:

Assumptions: Before going for the analysis, we assumed some things in order to get a better and authentic conclusion. For the present study we assume two things:

1. We have taken total expenditure as an approximation of income because data on income was not available.
2. Food expenditures given in various regions are summed up in order to come to total expenditure on food .
3. The direct expenditures (like fuel for cooking) incurred for food items have also been included under the purview of food expenditure.

Regression Analyses : The general forms of two main types of regression are the linear and multiple regressions. While the former takes into consideration only one independent variable, the latter consider more than one variables that we have taken as possession of land, social groups, education ,size of the household, types of fuel used by the households in the rural area ,religion, season etc. The form of both the regression can be written as:

Various functional forms that we have considered to follow for the desired results are:

1. Linear Function
2. Double-Log Function
3. Semi-Log Function
4. Working -Leser Function
5. Rank 3 Leser function

Elasticity of demand:

Consequently, formulation of Engel curves also calculated a good's income elasticity, which is roughly the percent change in *qi* that results from a one percent change in *y*, or formally

$\partial \log V / \partial \log(M)$

Goods with income elasticity's below zero, between zero and one, and above one are called inferior goods, necessities, and luxuries respectively, so by these definitions what Engel found is that food is a necessity. Elasticity's can themselves vary with income, so e.g. a good that is anecessity for the rich can be a luxury for the poor.

Uni-variate Model	Equations	Elasticity
Linear	$v_i = a_i + b_i M + \varepsilon_i$	$M/(a + bM)$
Double-log, Cobb Douglas(1927)	$\ln v_i = \ln a_i + b_i \ln M + \varepsilon_i$	b
Semi-Log	$\ln v_i = a_i + b_i M + \varepsilon_i$	bM
Working-Leser,Leser(1963)	$w_i = a_i + b_i \ln M + \varepsilon_i$	$1 + b/(a + b \ln M)$
Rank-3 Leser,Leser(1963)	$w_i = a_i + b_i \ln M + c_i/M + \varepsilon_i$	$(a + b \ln M + b)/(a + b \ln M + \frac{c}{M})$

Note:

$\boldsymbol{v_i}$ is the expenditure onfood **,**

$\boldsymbol{w_i}$ is the ratio of expenditure on food to total expenditure**,**

$\boldsymbol{M}$ is total expenditure,

$\boldsymbol{a, b \text{ and } c}$ Are coefficients for i^{th} commodity.

In our case $\boldsymbol{i = 1}$

Multivariate Analysis

The Eight input parameters that have been considered are:

- Household income
- Age
- Social Group: *scheduled tribe-1, scheduled caste-2, other backward class-3, others-9*
- Education level: *not literate, literate without formal schooling, literate but below primary, primary , middle , secondary , higher secondary , diploma/certificate course , graduate , postgraduate and above*
- Religion: *Hinduism, Islam, Christianity , Sikhism, Jainism, others*
- Household Type: *self-employed in non-agriculture, agricultural labour, other labour, self-employed in agriculture, others*
- Household size
- Whether any ceremony has been performed in last 30 days

Functional Forms of Independent Variables: We used 4 different functional forms of the variables in equations.

1. **Quantitative:** When the variable was quantifiable in nature, we

used its value in 3 possible ways: Direct, Inverse and Log .

2. **Qualitative**: For qualitative data like religion , we used binary notation of 1 and 0 for true and false respectively

Since we have taken all possible combinations of above forms, we had more than 50 variables to deal with. So to reduce the no. of variables, we ran a correlation analysis to prioritize them in order of their relevance and use the ones with highest correlation coefficient. Subsequently, we tried various combinations of the higher-ranking variables following backward elimination to arrive at the most optimal model.

IV. Results and major findings:

1. Univariate

Univariate Model	Sample1	Sample2
Linear	$v_i = 111 + .36M + \varepsilon_i$	$v_i = 116 + .36M + \varepsilon_i$
Double-log(Cobb-Douglas)	$ln\, v_i = .7 + .79\, b_i + \varepsilon_i$	$ln\, v_i = .9 + .76\, b_i + \varepsilon_i$
Semi-Log	$\ln v_i = 5.3 + .00077M + \varepsilon_i$	$\ln v_i = 5.4 + .00062M + \varepsilon_i$
Working-Leser,Leser(1963)	$w_{i\,=} 1.1 - .09 \ln M + \varepsilon_i$	$w_{i\,=} 1.2 - .10 \ln M + \varepsilon_i$
Rank-3 Leser,Leser(1963)	$w_{i\,=} 1.66 - .16 \ln M - 43/M + \varepsilon_i$	$w_{i\,=} 1.69 - .16 \ln M - 41/M + \varepsilon_i$

Source: Author's calculation

	SAMPLE-1	SAMPLE-2
	R-Square	**R-Square**
Linear	0.672654	0.708127
Double-log	0.868948	0.765843
Semi-log	0.59056	0.558684
Working-leser	0.178451	0.208089
Rank 3	0.18965	0.216279

Source: Author's calculation

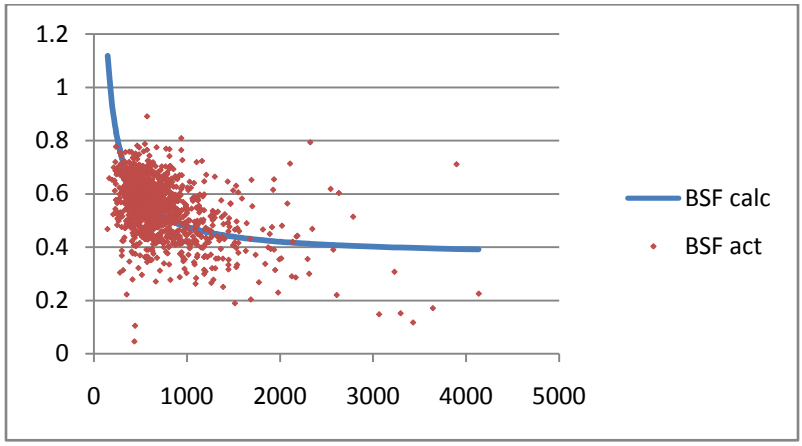
R-squared represents the proportion of the variation in the dependent variable explained. In the last two, the dependent variable is different, so their R-squared is not comparable with those of the first 3.

Model	Sample1	Sample2
Linear	$M/(111+.36M)$	$M/(116+.36M)$
Double-log	$.79$	$.76$
Semi-log	$.00077M$	$.00062M$
Working-leser	$1-.09/(1.1-.09l\,n\,M)$	$1-.1/(1.2-.1\,ln\,M)$
Rank-3	$(1.66-.16\ln M-.16)/(1.66-.16\ln M+\frac{43}{M})$	$(1.69-.16\,ln\,M-.16)/(1.69-.16\,ln\,M+\frac{41}{M})$

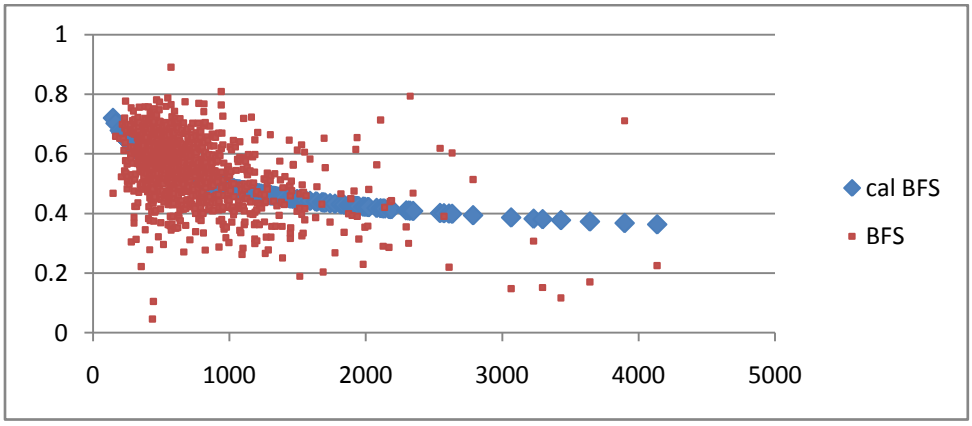
Source: Author's calculation

Engel Curves: Income vs Budget Share of Food Expenditure

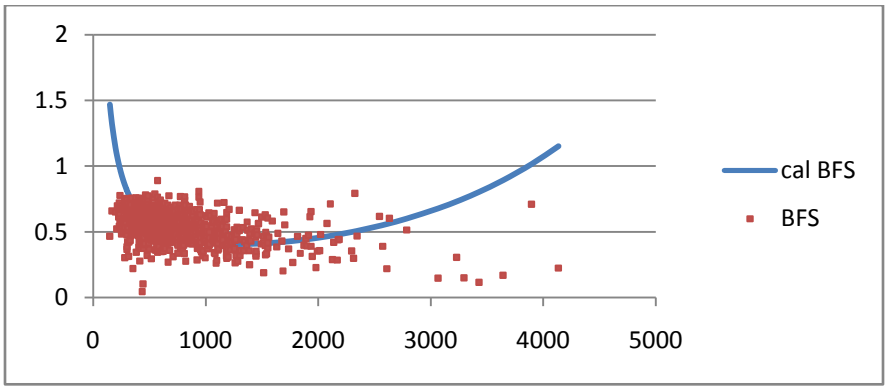
Linear Model:



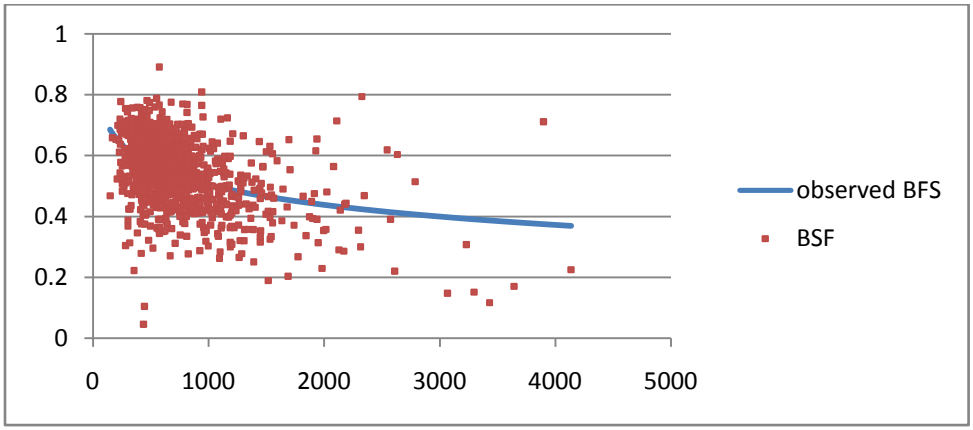
Double-log:



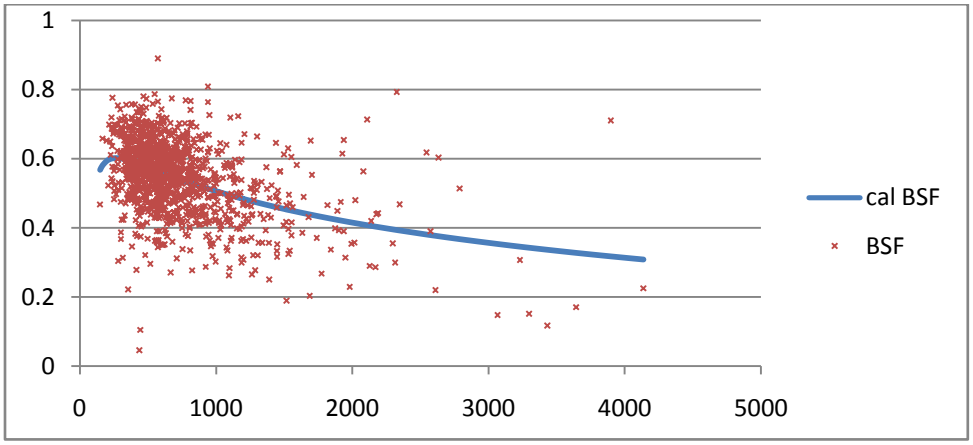
Semi-log:



Working-Leser Model:



Rank-3:



Source: Author's calculation and graph in excel

Multivariate analysis:

w_i = a (ln(M)) + b(1/M) + c(dEdu) + d (HH4) + e (MPD) + f (dSikh) + g (S3) + h (S2) + i (1/Size) + j (Gst) + k (1/Gst) + l (LN(Gst)) + m (PW) + n (1/Age)

Variable	SS1		SS2	
	Coefficients	*t Stat*	*Coefficients*	*t Stat*
Intercept	1.772	13.12	1.925	13.57
ln(M)	-0.195	-10.65	-0.215	-11.27
1/M	-54.451	-5.09	-58.007	-4.96
dEdu	-0.0058	-0.94	-0.0116	-1.87
HH4	0.0163	2.70	0.0059	0.97
MPD	0.027	3.16	0.028	3.21
dSikh	-0.077	-1.41	-0.092	-2.16
S3	0.022	3.19	0.029	4.23
S2	0.026	3.85	0.022	3.17
1/Size	0.051	2.56	0.103	5.57
Gst	0.0001	3.55	0.0002	4.63
1/Gst	0.062	3.01	0.028	1.41
LN(Gst)	-0.017	-2.27	-0.011	-1.54
PW	-0.0019	-3.21	-0.0019	-2.48
1/Age	-0.291	-1.59	-0.221	-1.17

Source: Author's calculation

M: Total expenditure, proxy for income
dEdu: Whether female member of family is educated
HH4: Whether self-employed in agriculture
MPD: Average meals per day for household
dSikh: Whether religion is Sikh
S3: Sub-round 3 dummy
S2: Sub-round 2 dummy
Size: No. of members In household
Gst: No. of meals served to guests in a month
PW: No. of days got work in public works over last year (avg for household)
Age: Average age of household

Note: dEdu, dSikh and Age do not pass the 95% confidence level test of statistical significance. However, in other combinations they did pass this test, so they have been retained in the final model.

Regression Statistics	*SS1*	*SS2*
Multiple R	0.498869	0.537197
R Square	0.24887	0.28858
Adjusted R Square	0.23971	0.279874
Standard Error	0.094467	0.09393
Observations	1163	1159

Source: Author's calculation

Discussion of results:

It is found that only a few of the assessed variables actually have any significant impact on the budget share of food. The findings are summarized in the following table:-

Variable	Impact on budget share of food
M	By and large, inverse relationship
dEdu	Lower budget share for higher education
HH4	Higher budget share for households self-employed in agriculture
MPD	Higher budget share for higher meals per day
dSikh	Lower budget share in the case of Sikhs
S3 & S2	Higher budget share during a particular half of the year
Size	Inverse relationship
Gst	By and large, increases with higher meals to guests
PW	Lower budget shares for greater employment in public works
Age	Increasing budget shares for higher average household age

Bibliography

[1.] C.E.V.Leser. (1963). Forms of Engel Functions. *Econometrica*, 694-703.

[2.] Chai, A. (2010). Restrospectives: Angel Curves. *Journal of Economics Perspectives*, 225-240.

[3.] Hikaru Hasegawa, H. K. (2001). Bayesian Analysis on Engel Curves: Estimation with measurement errors and an instrumental variable. *Journal of Business and Economic Statistics*, 292-298.

[4.] Deaton, A. and A. Tarozzi (2000), "Prices and Poverty in India", Princeton Research Program in Development Studies, Available from http:www.wws.princeton.edu/~rpds.

[5.] Foster, J., J.Greer, and E. Thorbecke (1984), "A Class of Decomposable Poverty Indices", Econometrica, Vol. 42.No. pp.

[6.] Government of India (1993), Perspective Planning Division, Planning Commission, Report of the Expert Group on Estimation of Proportion and Number of Poor.

[7.] Sukhatme, P.V. (1961), "The World's Hunger and the Future Needs of Food Supplies", Journal of the Royal Statistical Society, Series A, Vol 124, pp.463-525.

[8.] Sukhatme, P.V. (1965), Feeding India's Growing Millions, Asia Publishing House, Bombay.