Ind. Jn. of Agri.Econ.
Vol.63, No.3, July-Sept. 2008

An Enquiry into the Socio-Economic Status of Rainfed Communities – Logit Model Approach

T. Ponnarasi and K. Sita Devi*

I

INTRODUCTION

Rural poverty in Tamil Nadu is concentrated among those with marginal landholdings and dependent on rainfed agriculture. Dry land areas contribute about one half of India's production of coarse grain, cereals, pulses, oilseeds and cotton. Dryland agriculture is characterised by wide spatio-temporal variations in the productivity of crops due to uncertainty and high fluctuations in agro-climatic conditions. The poor among the farmers in the dryland communities suffer from double handicap. Firstly, they are located in a resource environment, which does not readily yield new and remunerative economic opportunities. Secondly, with a weak land base, it is unlikely that they would become viable by depending on agriculture alone. Thus, the households is rural areas are found to be in a low standard of living. Scheduled castes and tribes are highly represented among the poor. This is certainly due in part to their owning less land, and of lower quality, as well as other assets (particularly human capital), than households which are not of the scheduled castes. Important challenges in the non-income dimensions of poverty also remain. There are gender, caste, inter-district, and urban-rural disparities. The standard of living of a society, otherwise, said to be its well being and hence, its poverty which is a manifestation of insufficient well being depends on both monetary and non-monetary variables. Income as the sole indicator of standard of living is inappropriate and should be supplemented by other attributes or variables, e.g., housing, literacy, type of agricultural land possessed and so on. Hence, this paper attempts to study the socio-economic status of the rural households in rainfed areas. The specific objectives of the study are (i) To estimate the indices of levels of living of different types of households in rainfed area and (ii) To identify the factors influencing the households being poor.

*Lecturer and Reader, respectively, Department of Agricultural Economics, Annamalai University, Annamalai Nagar – 608 002 (Tamil Nadu).

II

DATA AND METHODOLOGY

Following a three-stage stratified random sampling method, 300 households from ten villages each of Kovilpatti block and Aruppukottai block from dry farming areas of Tamil Nadu were selected for the study. All the sample households were interviewed personally to collect the required primary data. The household enquiry included details on their socio-economic status including employment level, income and food consumption pattern, income spent on various food items, clothing, shelter, education, health, festivals, recreation, and other miscellaneous items, and also access to basic amenities like, safe drinking water, sanitation, school, transport, market facilities, communication and recreation facilities. The 'Z-test' analysis undertaken to find the homogeneity of sample mean indicated that the sample is homogeneous of the population. However, there existed high variations within the sample. Hence, the collected data were post-stratified into three categories of households, i.e., households with any irrigation source (borewell or open well)-Category I, households who practiced only dry farming (rainfed agriculture)-Category II and other worker households-Category III. Category I, II and III formed 27.00, 37.33, and 35.67 per cent of the sample households respectively. Average and percentage analyses were used. A comparative study of the households on their standard of living was attempted using "Composite Index of Standard of Living". Considering the major aspects of levels of living of the population an "Index of Deprivation (ID)" was also developed. This study utilised a logistic regression model to empirically quantify the relative influence of various factors influencing a household to be poor or non-poor.

III

ANALYTICAL FRAMEWORK

Composite Index of Standard of Living

Composite Index of Standard of Living was computed for each household combining the social and economic indicators using the scoring technique (Singh and Chand, 2000 and Puhazhendhi and Satyasai, 2000). The social indicators included the availability of electricity in the household, easy access to medical facilities, educational institutions, transport facilities, communication, recreation and market facilities, availability of proper sanitation within the house and access to safe drinking water. The economic indicators included the value of assets, income, consumption expenditure, savings and borrowings. The different indices were calculated as follows:

Index of social indicators of h-th household (S_h):

$$\sum S_i / \sum S_{i(max)}$$

Index of economic indicators of h-th household (E_h):

$$\sum E_j / \sum E_{j(max)}$$

Composite index of standard of living of h-th household ($CISL_h$):

$$w_1 S_h + w_2 E_h$$

where, S_i and E_j represent i-th social and j-th economic indicators, respectively. $S_{i(max)}$ and $E_{j(max)}$ are the maximum scores for i-th social indicator and j-th economic indicator. Weight w_1 is given by $\sum S_{i(max)} / \left(\sum S_{i(max)} + \sum E_{j(max)}\right)$ and w_2 is $(1-w_1)$.

Index of Deprivation (ID)

The indicators, which have shown significant difference between the poor and non-poor in their levels of living, were only considered in computing the deprivation Index. Various options were examined and the set of indicators (variables) is chosen in developing the Index of Deprivation (ID). The justification for selecting the above set of indicators is that the computed (total) targeting errors were found to be lower than in any other combination of indicators. However, alternative methods may be developed with a new/same set of characteristics or giving weights to the indicators considered in the present study. All the variables included in computing ID were given (i) equal weights and (ii) dichotomised as deprived (yes) or non-deprived (no) category (Singh, 2004).

For the identified socio-economic variables a score one is assigned if a particular household did not enjoy the social or economic benefit or status in the society. Otherwise a score zero is assigned. A simple Index of Deprivation (ID) is computed as the sum total of all such scores.

The composite index of deprivation for the 15 stated variables/indicators is computed for each individual household. Theoretically the ID value ranges between 0 and 15. If a household gets a value zero it indicates that the particular household/person has not been deprived in any one of the 15 aspects considered. On the other hand if it takes the value 15, that particular household is deprived in all aspects.

Factors Influencing the Households Being Poor

The logit model in this study postulates that P_i, the probability that a respondent i is poor, is a function of index variable Z_i summarising a set of the individual attributes. Hence, let us consider the following representation of a household being poor.

$$Pi = E(Y=1|X_i) = \frac{1}{1+e^{-(\beta_1+\beta_2 X_i)}} \quad \ldots.(1)$$

where, e is the familiar base of the natural logarithm. Now, let equation (1) be rewritten as

$$P_i = \frac{1}{1+e^{-z_i}} \quad \ldots.(2)$$

where $Z_i = \beta_1 + \beta_2 X_i$

Equation (2) represents the (cumulative) logistic distribution functions (Gujarathi, 1988).

It could be verified that as Z_i ranges from $-\infty$ to $+\infty$, P_i ranges between 0 and 1 and that P_i is non-linearly related to Z_i (i.e., X_i). However, we would encounter an estimation problem because P_i is not only non-linear in X but in the β's as well, as can be seen clearly from (1). This means that the familiar OLS procedure could not be made to estimate the parameters. But this problem is more apparent than real because (1) is intrinsically linear, which can be shown as follows:

If P_i, the probability of a household being poor is as given by (2), then, $(1-P_i)$, the probability of non-poor is

$$1 - P_i = \frac{1}{1+e^{-z_i}} \quad \ldots.(3)$$

Therefore, we can write

$$\frac{P_i}{1-P_i} = \frac{1+e^{z_i}}{1+e^{-z_i}} = e^{z_i} \quad \ldots.(4)$$

Now $\frac{P_i}{1-P_i}$ is simply the odds ratio in favour of poor – the ratio of the probability of poor to the probability of non-poor. Thus, if $P_i = 0.8$, it means that odds are 4 to 1 in favour of poor (Gujarathi, 1988).

Now, by taking the natural log of (4), we would obtain:

$$L_i = 1n\left(\frac{P_i}{1-P_i}\right) = Z_i$$

$$= \beta_1 + \beta_2 X_i \quad \ldots.(5)$$

that is, L, the log of the odds ratio, is not only linear in X, but (from the estimation viewpoint) linear in the parameters also. It might be noted that the linearity assumption of OLS does not require that the X variable be necessarily linear. So we can have X^2, X^3, etc., as regressors in the model. For our purpose, it is the linearity in the parameters that is crucial. L is called the logit, and hence the name logit model for (5).

Features of the Logit Model

1. As P goes from 0 to 1 (i.e., as Z varies from - ∞ to + ∞), the logit L goes from $-\infty$ to $+\infty$. That is, although the probabilities (of necessity) lie between 0 and 1, the logits are not so bounded.
2. Although L is linear in X, the probabilities themselves are not.
3. The interpretation of the logit model is as follows: β_2, the slope, measures the change in L for a unit change in X.

Estimation of the Logit Model

For estimation purposes, equation (5) can be written as follows:

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 X_i + u_i \qquad \ldots.(6)$$

To estimate the model, we need, apart from X_i, the values of the logit L_i. But now we run into some difficulties. If we have data on individual respondents, $P_i = 1$, if the respondent is poor and $P_i = 0$, if the respondent is non-poor. But, if we put these values directly into the logit L_i, we obtain:

$L_i = \ln\left(\frac{1}{0}\right)$ if the respondent is poor

$L_i = \ln\left(\frac{0}{1}\right)$ if the respondent is non-poor

Obviously, these expressions are meaningless. Therefore, if we have data at the micro or individual level, we cannot estimate (equation 6) by the standard OLS routine. In this situation, one may have to resort to the maximum likelihood method to estimate the parameters (Uma Devi and Prasad, 2006).

Within the logit framework discussed above, this study has postulated that the probability of an individual being poor (L_i) is dependent upon the attributes like age, percentage of literates, category (land holding), social status, percentage of earners in the household, household income, man-days employed.

The index variable P_i indicating whether the respondent is poor or non-poor has been expressed as a linear function of the independent variables. Thus, the logit regression model has been specified as follows.

$$L_i = \alpha_i + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + u_i$$

where,

X_1 - Age of the respondents, in years,
X_2 - Percentage of literates in the household,
X_3 - Category of the respondents, 1 if category-II, 0, otherwise,
X_4 - Social status of the households, 1 if scheduled caste, 0, otherwise,
X_5 - Percentage of earners in the household,
X_6 - Annual income of the household, in rupees,
X_7 - man-days of employment of the household,
β_i's - Parameters to be estimated,
u_i - Error term.

IV

RESULTS AND DISCUSSION

As stated in the objective of this study the standard of living of the sample households was analysed using an aggregate measure encompassing social as well as economic aspects. The composite index of standard of living has been worked out by assigning scores to the identified economic variables and social variables, index of economic indicators and index of social indicators were also analysed separately for each of the three categories of households.

It could be seen from Table 1 that the estimated average value of index of standard of living was lowest in category II households, which was 34, whereas it was higher in category I households, followed by category III which accounted to 62 and 46, respectively. Also, the average value of index of economic indicator showed a similar trend. However, the average index of social factors was high in the case of category I households (50), followed by category II households (48) and it was lowest in households of category III.

The distribution of households according to the value of composite index clearly indicated that about 50 per cent of the category I households were found to be distributed in the index value of above 60 and 38.27 per cent of them were distributed in an index value ranging between 40 and 60. In case of the category III households, an almost similar situation obtained, i.e., around 45 per cent of them had been placed in the index of above 40, and about 50 per cent of them were distributed in the index range of 20-40. However, in the case of category II households, only 16 per cent of the sample households were distributed in the index value ranging between 40 and 80. A good majority of about 80 per cent of these households were having a lower index of standard of living of 20-40.

With regard to the index of economic indicator also only 11 per cent of category II households were distributed in the index value of 40-80. However, around 63 per cent of these households were distributed in the index value of between 40 and 80

TABLE 1. COMPOSITE INDEX OF STANDARD OF LIVING FOR THE SAMPLE HOUSEHOLDS

(per cent)

	Social Index			Economic Index			Composite Index		
Index (1)	Cultivators (2)	Agricultural labourers (3)	Other workers (4)	Cultivators (5)	Agricultural labourers (6)	Other workers (7)	Cultivators (8)	Agricultural labourers (9)	Other workers (10)
Upto 20	-	-	-	-	11.61	8.41	-	3.57	3.74
21-40	35.80	35.72	47.66	12.35	76.79	47.66	11.11	79.46	49.54
41-60	37.04	47.32	32.71	28.39	8.03	17.76	38.27	14.29	23.36
61-80	27.16	16.96	19.63	33.33	3.57	14.95	40.74	2.68	20.56
81-100	-	-	-	25.93	-	11.21	9.88	-	2.80
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Average Index	50	48	45	65	30	46	62	34	46

with respect to the index of social indicator. It could also be noted that around 59 per cent of the category I households were distributed in the index value of between 60 and above and the remaining households lay below the index of 60. The category III households were more pronounced in the economic aspects than the social aspects. Thus it could be concluded that the category II households were found to be the disadvantaged category, whose standard of living was lower as compared to the other two categories of the sample households in both economic and social aspects.

Index of Deprivation (ID)

The index of deprivation for the selected non-monetary/social variables had been computed for each of the individual household. Theoretically, the ID value ranges between 0 and 15. The percentage distribution of households by the level of deprivation categorised as not deprived, less deprived, moderately deprived and the most deprived is presented in Table 2.

TABLE 2. DISTRIBUTION OF HOUSEHOLDS BY LEVEL OF DEPRIVATION

(Numbers)

Sl. No. (1)	ID (2)	Category I (3)	Category II (3)	Category III (4)	All Samples (5)
1.	Not deprived (0-3)	65 (80.25)	-	38 (35.51)	103 (34.33)
2.	Less deprived (4-7)	16 (19.75)	63 (56.25)	59 (55.14)	138 (46.00)
3.	Moderately deprived (8-11)	-	49 (43.75)	10 (9.35)	59 (19.67)
4.	Most deprived (12-15)	-	-	-	-
	Total	81 (100.00)	112 (100.00)	107 (100.00)	300 (100.00)

Note: Figures in parentheses represent percentages to respective total.

It could be seen from Table 2 that around 80 per cent of the category I households lay in a not deprived state and the remaining households were less deprived. The less deprived households had an ID value ranging between 4 and 7, which accounted to 56.25 per cent of category II households and 55.14 per cent of category III households. Also, around 44 per cent and 9 per cent of the category II and category III households, respectively were found to be moderately deprived with ID values ranging between 8 and 11. It is also seen that 19.67 per cent of the total sample households were moderately deprived of the selected social indicators and 46 per cent of them were less deprived, where as 34.33 per cent of the total households were not deprived. However, a state of deprivation for the selected social indicators does persist among the sample households. Thus, it could be inferred that the sample households were found to be deprived based on the social indicators, economic

indicators and housing indicators. In sum the category II households were more deprived than the other two categories of sample households.

Factors Influencing a Household Being Poor

The logit framework has postulated that the probability of a household being poor was dependent on the socio-economic characteristics of the households. The Maximum-Likelihood Estimate of the coefficients of the logit model for the respondents is presented in Table 3.

TABLE 3. MLE COEFFICIENTS FOR LOGIT MODEL

Sl. No. (1)	Variables (2)	Logit MLE coefficient (3)	Standard error (4)
	Intercept	4.1547***	1.3505
1.	Age	0.0019***	0.0059
2.	Percentage of literates	-0.0857*	0.0534
3.	Category	-0.9344*	0.4854
4.	Social status	0.3958	0.4383
5.	Percentage of earners	-0.3971*	0.2150
6.	Income	-0.0484**	0.0234
7.	Man-days of employment	-0.0027*	0.0015
	Count R^2	0.87	
	Number of observations	300	

*, ** and *** Significant at 10, 5 and 1 per cent level, respectively.

The results show that the specified logit model was significant at ten per cent level of probability. The level of Count R^2 obtained was 0.87, which indicated the good predictive ability of the model. The estimation yielded the expected signs for the coefficients of all the independent variables except social status. The results indicated that literacy percentage, category, man-days of employment, percentage of earners in the household and income of the household were negative and significant. Thus, it could be inferred that one unit change in the negative and significant slope coefficients would decrease the probability of the respondent being poor by their appropriate percentages. The coefficient of the independent variable age is positive and significant and indicated that the change in age would increase the probability of the respondent to be poor. The coefficient of the independent variable social status was positive indicating that the probability of SC/ST respondents to be poor and non-SC/ST respondents to be non-poor. However, this coefficient is not significant, and hence the social status of the respondent could not influence their probability of being poor. Also, the case is true among the sample households. The non-SC/ST households were also found to be poor. The results of this analysis would imply that the probability of a respondent being poor would be influenced by the factors/variables considered in this model except that of the social status of the respondent.

V

CONCLUSIONS AND POLICY IMPLICATIONS

Composite index of standard of living was estimated, and the index value was lowest in category II households, which was 34, whereas it was higher in category I households, followed by category III households which accounted to 62 and 46, respectively. The category II households were found to be the disadvantaged category, whose standard of living was lower as compared to the other two categories of the sample households in both economic and social aspects.

Index of Deprivation (ID) was computed using the scoring technique for the identified 15 non-monetary indicators. A state of deprivation for the selected social indicators does persist among the sample households. Thus, it could be inferred that the sample households were found to be deprived based on the social indicators, economic indicators and housing indicators. However, the category II households were more deprived than the other two categories of sample households.

Logistic regression model adopted to study the factors influencing a household to be poor, showed that the level of Count R^2 was 0.87, which indicated the good predictive ability of the model. The estimation yielded the expected signs for the coefficients of all the independent variables except social status.

Since the households, with dry land faming are found to be more deprived and poor the planners could encourage the establishment of employment generating activities in rainfed areas through diversified farming enterprises. The levels of living of rural sector was found to be very low especially among the rainfed farmers due to the lack of rural infrastructure. Hence, government might redouble the efforts to strengthen the rural infrastructure facilities through various welfare schemes.

REFERENCES

Gujarathi, Damodaran N. (1988), *Basic Econometrics,* Mc Graw-Hill Company, New Delhi.

Umadevi, K. and Y.E. Prasad (2006), "A Logistic Regression of Risk Factors for Disease Occurrence on Coastal Andhra Shrimp Farms", *Indian Journal of Agricultural Economics,* Vol. 61, No.1, January-March, pp. 123-133.

Puhazhendhi, V. and K.J.S. Satyasai (2000), *Micro Finance for Rural People –An Impact Evaluation,* National Bank for Agricultural and Rural Development, Mumbai.

Richa, Singh (2004), "Rural Infrastructure, Agricultural Development and Poverty in India: An Inter-State Study", *Journal of Rural Development,* Vol.23, No.1, pp. 31-57.