ISSN 1023-1072

Pak. J. Agri., Agril. Engg., Vet. Sci., 2016, 32 (2): 258-267

FACTORS TRIGGERING POVERTY IN DISTRICT THARPARKAR SINDH, PAKISTAN

V. Suthar[1], M. S. Sarki[2], N. A. Qureshi[1], B. K. Solangi[3] and T. Mangan[4]

[1]Department of Statistics, [2]Department of Soil Sciences,
[3]Department of Entomology, [4]Department of Agricultural Economics,
Sindh Agriculture University, Tandojam, Pakistan

ABSTRACT

Poverty is one of the most unadorned problems of Pakistan; about over one third of the population is declared as poor. Several programs planned by the government and NGOs to eliminate the poverty could not succeed in their high honored goals. Current study has strived to discover the factors related with poverty in Tharparkar district of Sindh, Pakistan. The survey data in the study showed that three-fourth of the population this district lives below poverty line. Maximum likelihood method was used to estimate the factors associated with poverty. On an overall basis the logistic regression model could correctly classified about 81.5% of the total cases. Furthermore, the highest probabilities for poor were likely for less educated household heads, bigger household size and the female household heads were more exposed than male household heads to risk of being poor at the level of significance set at 0.05.

Keywords: maximum likelihood method, poverty, survey data, Tharparkar-Sindh

INTRODUCTION

Poverty in Pakistan is an increasing social problem and represents the critical challenge to be addressed. It is estimated that about 32 percent of Pakistan's population is below food poverty line raised from a level of twenty six percent in 1988 (Government of Pakistan, 2002); and about forty-four percent are below the poverty line on the human poverty index (United Nations Development Report, 2002). District Tharparkar is the desert of Sindh, which is ranked as world's seventh largest and hottest desert. Tharparkar lies in the southeastern part of Sindh province and is one of the most deprived regions of Pakistan. It has about one million population. Thar desert depends on rainfall, where agriculture and postural activity is the main source of income of the people. After the rains, the desert subsoil aquifers are recharged and the pasture lands are regenerated. However, by February, the aquifers often depleted and the pasture lands dry up. Households identified the lack of water as the most severe community problem in the area. This has resulted in increased workload for women and children in

Corresponding author: vsutahar@yahoo.co.uk

Tharparkar. Many women are covering longer distances to fetch water, as the underground water sources have receded and surface sources have dried up. In many areas, there is increased crowding at the water sources due to fewer and lower available water resources (ADBP, 2002).

The district lacks in basic human needs, such as adequate and nutritious food, clothing, housing, drinking water, and also lacks in opportunities like health services, education facilities, transportation, sanitation, source of communication, etc. One can rename the district Tharparkar as "Poverty". There is no single facility of life available in this district, therefore, it is called extreme/ absolute poverty based region (Suthar, 2004). As the water is too brackish for vegetation, vegetables in this area are not grown. In this study, it is important to measure the extent of influence of each factor associated with poverty.

The purpose of the study was to measure the poverty and to investigate the socio-economic conditions of Tharparkar district of Sindh-Pakistan. Moreover, this study provides a baseline for policy makers to put focus on factors responsible for increasing poverty among people, so that a poverty alleviation plan can be framed in this backward district of Sindh.

METHODOLOGY

The main objective of this study was to explore the factors that significantly contributed in enhancing the extreme poverty through applying more advanced and sophisticated statistical technique such as logistic regression model to sample data of households collected from district Tharparkar of Sindh Province. Variables were created, labeled and categorized in Statistical Package for Social Sciences (SPSS, 2009) Version 17). Brief methodology about the study conducted is elaborated here in:

Population, sample and sampling method

The estimates of this study reflect district Tharparkar of Sindh province of Pakistan. However, due to limited resources, a sample of 325 respondents was justified referring the table of determining sample size Wunsch (1986) at $\pm 5.5\%$ error rates. According to Wunsch, used when population is greater than 3000, the population is considered infinity and a fixed sample sizes regardless of population size are recommended against different error rates at 0.05 of significance level. Three-stage cluster sampling was used to select 325 heads of households from all 4 talukas/ tahsils of district Tharparkar. In first stage, two union councils were selected from each taluka; in second stage, five villages from each randomly selected union council were selected; in third stage, a random sample of 8 respondents from each village was randomly selected. Thus, a total of 325 household heads were selected from 8 marked union councils. A total of 314 questionnaires were administered out of 325 due to non-availability of 11 respondents at their residences.

Poverty criterion

The Bureau of Statistics of Pakistan (2004) conducted a sample survey which after taking into consideration the inflation in the country, set the poverty line at Rs. 848 per capita income per month (Khan, 2005). Garza (2003) has also used the same criteria to define the poverty line. Therefore, first of all per capita

income of respondents was calculated and then those having more than per capita income per month more of Rs. 848, were categorized as “not poor” while those having less than this value were termed as poor.

Statistical analysis

Logistic Regression Model (LRM) is part of a category of statistical models called generalized linear models. This model allows one to predict a discrete outcome, as group membership, generally, the dependent or response variable is dichotomous, such as success/ failure. Thus, in instances where the independent variables are categorical, or a mix of continuous and categorical, logistic regression analysis is preferred. The logistic regression analysis is used to model Pr (y=1 | x) = F (xβ). To do this we will make use of the logit transformation.

Let P = Pr (y=1 | x), which can also be written as:

π (x) = E (y|x)

Let odds = P/1-P = $\frac{\pi(x)}{1-\pi(x)}$

Let log odds or logit g (x) = ln (P/(1-P)) = $\ln\left[\frac{\pi(x)}{1-\pi(x)}\right]$

The coefficients for logistic regression are estimated using maximum likelihood method using iterative procedure. This is because the effects in logistic regression model the solutions are nonlinear in β_0 and β_1, the goal is to find the coefficients that make the data most likely. This is done by the likelihood function.

The odds would be exp (xb) = exp ($\beta_0 + \beta_1 x$) which can be rewritten as exp (β_0) exp ($\beta_1 x$). If we increase x by one we get exp ($\beta_0 + \beta_1 (x+1)$) = exp ($\beta_0 + \beta_1 x + \beta_1$) which, in turn, can be rewritten in logarithmic form as exp (β_0) exp ($\beta_1 x$) exp (β_1). Next, to compare the odds before and after adding one to x, we compute the odds ratio,

$$\frac{\exp(\beta_0)\exp(\beta_1 x)\exp(\beta_1)}{\exp(\beta_0)\exp(\beta_1 x)} = \exp(\beta_1),$$

i.e., the odds ratio for a one unit change is just the exponential log odds coefficient.

RESULTS AND DISCUSSION

Primary data were collected from socio-economic survey. Analysis was performed using descriptive and inferential statistics. Mean, standard deviation, percentages and proportions were calculated in descriptive statistics, while in inferential statistics logistic regression was applied to explore the factors associated with the extreme poverty.

Descriptive statistics are summarized in Table 1. The average age of respondents was 45 years with the standard deviation of 12.83 years. The age estimate exactly coincides with the age estimates provided by Jaffri (1999) and Garza (2003), who reported that the average age was 44.5 and 44 years, respectively. The average income of household head per month of respondent was Rs. 3362 and average of total income per month of the household was Rs. 4047 while the average of total expenditure per month was Rs. 3343 per household with minimum and maximum of Rs. 550 and Rs. 12650, respectively. Average food expenditure and non-food expenditure were estimated as Rs. 2088 and Rs. 1255, respectively. Average income per capita was calculated to be Rs. 597. The average household size was 7.47 while its minimum and maximum were recorded to be 2 and 18, respectively. These estimates are also in line with research estimates of Jaffri (1999) and in partial agreement with estimates by Arif (1999) who reported that the average household size was 7.5 and 8.7, respectively. The average number of earners was 2.0 and average of male members was 3.45 per household. The average number of female members, number of dependent children, number of students, and number of age dependency per household were 3.90, 2.40, 1.38, and 0.48, respectively, while the wage rate in the vicinity was estimated to be Rs. 53. The table also presents standard error, standard deviation and skewness of all these variables.

Table 1. Descriptive statistics of socio-economic key indicators

Variable	Minimum	Maximum	Mean		SD
			Statistic	S.E.	
Age	18	85	44.84	.72	12.83
Income of HHH/m	400	15500	3361.91	125.92	2231.28
Household size	2	18	7.47	.17	3.05
Number of earners	1	5	2.00	0.059	1.06
Male members	1	10	3.45	.11	1.98
Female members	0	9	3.90	.11	1.98
No. of child dependency	0	8	2.40	.10	1.85
No. of student	0	5	1.38	.00663	1.17
No. of aged dependency	0	3	.48	.0039	.70
Income of spouse	0	5700	231.39	31.43	556.98
Wage rate in vicinity	25	100	53.07	.63	11.13
Food exp. HH / m	200	6000	2087.94	57.93	1026.55
Non-Food exp. HH/m	20	7250	1254.94	52.83	936.20
Health facility	1	3	1.53	.00499	.88
Total Income per HH	400.00	18500.00	4047.18	159.79	2831.51
Income per capita	83.33	3400.00	597.01	26.27	465.51
Total expenditure	550.00	12650.00	3342.88	95.18	1686.56

Logistic regression model was estimated on 314 cases; out of these, 233 were recoded to be poor having income per capita less than Rs. 848 per month to assess prediction of group membership in one of two categories of outcome (poor and not poor) on the basis of variables that were assumed to enhance the

extreme poverty. About 74.2% households were recorded to live below poverty line while only 81(25.80%) were reported to live above poverty line; in other words, three-fourth of the households surveyed was in extreme poverty. These results greatly coincide with that of Mahmood *et al.* (1989) using HIES-1984/85 rural and urban data and adjusting for family size and composition. It is observed that 76% households in urban and 40% in rural areas were poor, where intensity of food poverty was very high and was more pronounced in rural compared to urban areas. These results also coincide with that of TRDP (2005) report, which showed that about 60 percent of the households live in poverty in the Tharparkar district of Sindh province of Pakistan.

Table 2. Categorical variables parameter coding for occupation variable

Category	Frequency	Percent	CF*	Parameter coding						
				(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tenant	10	3.2	3.2	.000	.000	.000	.000	.000	.000	.000
Peasant	23	7.3	10.5	1.00	.000	.000	.000	.000	.000	.000
Agri. Labor	120	38.2	48.7	.000	1.00	.000	.000	.000	.000	.000
Landlord	44	14.0	62.7	.000	.000	1.000	.000	.000	.000	.000
Businessmen	46	14.6	77.4	.000	.000	.000	1.00	.000	.000	.000
Herder	11	3.5	80.9	.000	.000	.000	.000	1.00	.000	.000
Artisan	22	7.0	87.9	.000	.000	.000	.000	.000	1.00	.000
Any Other	38	12.1	100.	.000	.000	.000	.000	.000	.000	1.00

CF* = Cumulative Frequency

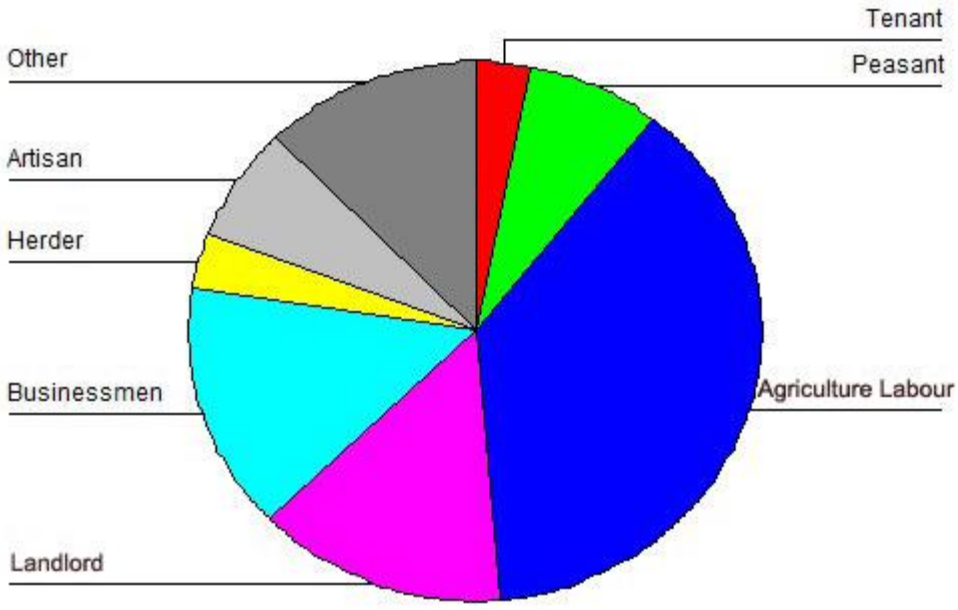

Figure 1. Pie chart by categories of occupation variable

Table 2 shows that there were 10 (3.2%) tenants, 23(7.3%) peasants, 120 (38.2%) agricultural labour, 44(14.0%) landlord, 46 (14.6%) businessmen, 11 (3.5%) herder, 22 (7.0%) artisan and 38 (12.1%) have different occupations designed as "others", Figure 1 also presents the clear vision of each occupational categories. The findings are similar to that of Jaffri (1999) who found that the highest poverty in labors and in agriculture workers where lower level of poverty was seen in managerial and professionals workers. It was observed that the majority of the respondents were agricultural labour in this study. Coding scheme for logistic regression analysis to determine the category-wise behavior of occupation by taking tenants as reference category (setting all the dummy variables at zero) is elaborated in Table 2. Caudill (1987) and Oksanen (1986) explored the use of observation-specific dummy variable in linear probability and logit models.

Testing hypothesis about the coefficients

The following hypotheses were formulated to explore the factor(s) that significantly contributing to enhance poverty in the region:

H_0:None of the factor(s) significantly enhances the poverty [$\beta_1 = \beta_2 = \beta_3 = ... = \beta_P = 0$]

H_A: At least one factor significantly enhances the poverty [At least one $\beta \neq 0$]

Estimating the parameters and assessing the goodness-of-fit of the predictive model

Table 3 illustrates that the coefficient for education was -0.569 and its estimated standard error was 0.17, shown in the column labeled as S.E. The Wald Statistic was $(\beta/se)^2$ = (-0.473/ 0.162) $^2 \cong$ 11.197. The significance level for Wald Statistic is showing in the column labeled as *Sig***.** It is obvious that all the coefficients such as sex, education, occupation, number of earners, household size, age dependency, wage rate and metallic road to be significantly different from 0 at the level of significance set at 0.05.

Likelihood is the probability of observed results; given the parameter estimates and -2 times the log likelihood (-2LL), is a measure of how well the estimated model fits the categorical data. For the logistic regression model that contains only the constant, -2LL is 358.534, as shown in Table 4. Goodness-of-Fit statistic for the model with significant independent variables, the value of -2LL for the proposed model is 217.447, which is smaller than the -2LL for the model containing only a constant. The *Cox and Snell* R^2 and *Nagelkerke* $\tilde{R}^2$ are statistics that attempt to quantify the proportion of explained "variation" in the logistic regression model. The value of *Cox and Snell* R^2 is 0.312 reveals that about 31.2% of the variation in the dependent variable is explained by the fitted model. But the problem with this measure for logistic regression is that it cannot achieve a maximum value of 1. Likewise, the *Nagelkerke* $\tilde{R}^2$ is shown in the

same table as 0.458, which indicates that about 45.8% of the variation in the outcome variable is explained by the logistic regression model.

Table 3. Parameter estimates for logistic regression model of significant independent variables (variables in the equation)

Variable	B	S.E.	Wald	df	Sig.	R
Sex	2.784	1.385	4.037	1	0.045*	0.075376
Education	-.569	0.170	11.197	1	0.001*	-0.160161
Migration Status	1.136	0.676	2.824	1	0.093	0.04794
Occupation			37.683	7	0.000*	-0.3155
Peasant	-.407	1.119	0.132	1	0.716	0.000
Agricultural Labour	1.331	1.084	1.505	1	0.220	0.000
Landlord	-.967	1.030	0.882	1	0.348	0.000
Businessman	-2.440	1.035	5.561	1	0.018*	0.0997
Herder	-1.313	1.297	1.025	1	0.311	0.000
Artisan	-.749	1.100	0.464	1	0.496	0.000
Other	-1.513	1.044	2.098	1	0.148	-0.016533
Household Size	.477	0.088	29.569	1	0.000*	0.277297
Number of Earner	-.686	0.200	11.787	1	0.001*	-0.165219
Age Dependency	.562	0.288	3.808	1	0.051	0.071012
Wage Rate	-.033	0.014	5.122	1	0.024*	-0.093315
Metallic Road	-2.045	0.697	8.612	1	0.003*	-0.1358
Constant	-.515	1.881	.075	1	.784	

*Significant at 95% C.I.

There is another additional chi-square entry in the table, labeled as *Chi-square*. The model chi-square is the difference between –2LL for the model with only a constant and –2LL for the current (full) model. The model chi-square tests the null hypothesis that the coefficients for significant terms in the current model, except the constant, are 0. The chi-square of the proposed model was calculated as 141.087 (difference between minus two log likelihood estimates). The degree of freedom for the model chi-square is the difference between the numbers of parameters in the two models i.e., 15. The model chi-square is 141.087, which is highly significant at 1% level of significance that indicates that the model fits well to the survey data with the significant variables in the proposed model.

Table 4. Statistics for model containing only constant and I.Vs and model Chi-Square

2 Log Likelihood (with only constant)			358.534
-2 Log Likelihood (proposed model)			217.448
Cox and Snell R^2			0.312
Nagelkerke $\widetilde{R}^2$			0.458
Model	Chi-square	df	Sig.
	141.087	15	0.0000

*Independent Variables (I.Vs)

Classification of cases in the study

One method of assessing the success of a model is to evaluate its ability to predict correctly the outcome category of cases for which outcome is known. If a case has diagnosed diseased, for instance, it can be seen if the case is correctly classified as diseased on the basis of other predictor variables. Table 5 shows the classification of cases of 314 individuals. It is obvious from the data that 218 households were in extreme poverty and they were correctly predicted by the model. Likewise, 38 individuals are not in poverty and the model correctly predicted them. On an overall basis the proposed model could correctly classified about 81.5% of the total cases. The proposed model was correctly predicted about 51% of total cases, while the logistic model of the present study correctly classified about 81.5% of cases.

Unlike diagonal entries, the off-diagonal entries indicate the incorrectly classified individuals by the model. There were 15 individuals below the poverty line, but they were incorrectly predicted “not poor” by the model. Likewise, forty-three cases were incorrectly classified as poor by the model while in real situation they were recorded to be not poor. On an overall basis, there were 18.5% of the cases that were incorrectly classified by the proposed model.

Table 5. Classification of households for logistic regression model by entered method

Classification		**Predicted**		**Percentage**
		Not Poor (0)	**Poor (1)**	
Observed	Not poor (0)	38	43	46.9%
	Poor (1)	15	218	93.6%
Overall				81.5%

CONCLUSION

A study was conducted in district Tharparkar to investigate the factors responsible for poverty using logistic regression. A sample of 325 households was decided on ±5.5 error rate and 95% confidence interval. Out of 325, 314 household heads were interviewed. A questionnaire was developed after intensive review of the literature and in consultation with researchers. The survey data in the study showed that three-fourth of population of this district lives in poverty. Maximum likelihood method was applied to estimate the factors associated with poverty. The highest probabilities for poor were likely for less educated household heads, bigger household size, seasonally migratory families, and low wage rates in vicinities. The female household heads were more exposed than male household heads to risk of being poor. The government and NGOs are liable to eradicate the poverty in the study area and launch several programs addressing the poverty related factors.

RECOMMENDATIONS

The estimates of the study appeal to the policy makers to launch a multidimensional poverty alleviation program in Tharparkar district where about three-fourth of the population lives below poverty line i.e. Rs.848 while at national

level about one-third of the population lives in poverty. If criterion, i.e., $ 1 per day, of the poverty suggested by international agencies is applied, almost all the population will fall below the poverty line. In the light of the factors found to be significant, the following recommendations are made:

- Program for female capacity building for example embroidery, cloth sewing and animal rearing be launched since female household heads were more exposed to poverty in comparison of their male counterparts.
- Educational programs should be launched because poverty was negatively related with the education level of the household heads.
- Special allowances like social security in developed countries should be allowed to elderly people of Tharparkar district.
- Infrastructure should be developed in the area which includes electricity, wells, schools, and hospitals, by taking metallic roads on the top priority. After construction of the metallic roads, development process will increase manifold, which will help to reduce poverty menace.

REFERENCES

ADBP. 2002. Poverty in Pakistan: issues, causes, and institutional responses, Pakistan Resident Mission, Islamabad.

Arif, G. M. 1999. Measuring poverty in Pakistan: A Critical review of recent poverty lines in human condition report. Centre for Research on Poverty Reduction and Income Distribution (CRPRID) Islamabad.

Caudill, S. B. 1987. Dichotomous choice models and dummies variables. The Statistician, 35: 381-383.

Bureau of Statistics of Pakistan. 2004. Labour Force Survey 2003-2004, 24th Issue, Statistics Division, Government of Pakistan, Islamabad.

Franekel, J. R. and N. E. Wallen. 1993. How to design and evolutes research in education, New York: McGraw, Hill Inc.

Garza, R. J. 2003. The Determinants of poverty in Mexico: 1996. Ph.D. Diss., University of Missouri-Columbia, 2000.

Government of Pakistan. 2002. Economic Survey of Pakistan, Finance Division, Islamabad.

Hosmer, D. W. and S. Lemeshow. 2000. Applied logistic regression. John Wiley and Sons, New York.

Jaffri, S. M. 1999. Assessing poverty in Pakistan in a profile of poverty in Pakistan. Centre for Human Development and United Nation Development Program, Islamabad.

Khan, A. A. 2005. The poverty debate: Dawn Business and Economic Review, Comment and Analysis, Dawn News Paper (http://www.dawn.com, 03-01-05) of the Pakistan.

Mahmood, S., H. S. Khalid and T. Mahmood. 1989. The food poverty in Pakistan. The Pakistan Develop. Review, 27 (4): 11-29.

Oksanen, E. H. 1986. A Note on observation-specific dummies and logit Analysis. The Statistician, 35: 413-416.

SPSS. 2009. Base Application Guide. USA: Inc., 444 North Michigan Avenue, Chicago, IL, 6061.

Suthar, V. 2004. A Logistic Regression Model for the Identification of factors Associated With Poverty in Tharparkar district of Sindh-Pakistan. Master's Thesis.

TRDP. 2005. Household Economy Assessment Report: Thar desert livelihood zone, Tharparkar district, Thar deep Rural Development Program, Mithi, Sindh Province, Pakistan.

United Nations Development Report. 2002. Human Development Report, Washington, DC.USA.

Wunsch, D. R. 1986. Forum Feature: Action research in business education. Business Education Forum, 5: 31-34.

(Accepted August 04, 2016)