Available online at www.isas.org.in/jisas
Journal of the Indian Society of Agricultural Statistics 73(2) 2019 137–142

A Study on Factors Affecting Rural Poverty in Odisha

Bigyana Nanda Mohanty[1], A.K.P.C. Swain[2] and Firoz Khan[1]
[1]Directorate of Economics and Statistics, Odisha, Bhubaneswar
[2]Utkal University, Bhubaneswar

Received 20 October 2016; Revised 27 March 2019; Accepted 26 April 2019

SUMMARY

The purpose of this study is to identify the most important factors affecting poverty in rural Odisha by using household level 68th round primary data (2011-12) collected by the National Sample Survey Office (NSSO) on consumer expenditure. In this study a logistic regression analysis is undertaken where the dependent variable is a dichotomous variable (Y), coded as Y = 1 for the household below official poverty line and Y = 0 for the household above poverty line and thirteen explanatory variables are age and sex of the head of the household, education, marital status of the head of house hold, family size, at least one salary earner in the household, social group of the household, total land possessed by the household, total cultivated land of the household, number of dwelling units in the household, source of energy for lighting and cooking and percentage of monthly per capita expenditure under food items. It was found that age, marital status and general education level of the head of the household, family size, at least one salary earner in the household, social group of the household, source of energy for lighting and percentage of monthly per capita expenditure under food items are significant predictors affecting poverty in rural Odisha.

Keywords: Poverty, Logistic regression, National Sample Survey Office, Odisha.

1. INTRODUCTION

Poverty is a world-wide phenomenon and also exists in different countries in different forms and in varying degrees with respect to its extended definitions of denial of choice and opportunities. It has since long attracted attention of eminent sociologists, economists and above all the planners for proactive planning and policies to address the poverty issues. Poverty is the outcome of various interactive socio-economic factors, which need to be identified so as to help the planners to formulate policies and welfare programmes to uplift the socio-economic status of the poor.

The State of Odisha is the 10th largest State in the Indian Union situated on the Eastern Coast of India with geographical area of 155707 square kilometer having 3.47%. India's population as per 2011 population Census. About 85% population of Odisha live in rural areas. According to 2011 Census, the Scheduled Castes (SC) and the Scheduled Tribes (ST) constitute 17.1% and 22.8% of the total population of Odisha respectively. In spite of rich cultural heritage, and being endowed with abundant natural resources and favourable political and social climate, Odisha stands at the bottom of the economic development as compared to most of the States of the Country.

As observed by the Planning Commission, Government of India, Odisha is one of the poorest States in India having high incidence of poverty 32.59% during 2011-12. The rural picture of State's poverty was slipped down to 35.7% during 2011-12 from 57.6% during 1987-88. But still, Odisha remains well above the national average of poverty. According to various reports of the Planning Commission, Government of India the percentage of poor and rural poor living in Odisha compared to India as a whole from 1973-74 to 2011-12 is presented in Table 1.

The visible downward movement in reduction of poverty was noticed (Table 1) in the State from the year 1987-88(55.6%) to 2011-12(32.6%) corresponding to 38.4% to 21.9% at all India level respectively.

Corresponding author: Bigyana Nanda Mohanty
E-mail address: bigyanmohanty65@yahoo.com

Table 1. Percentage of Poor in Odisha and India from 1973-74 to 2011-12

Year	Odisha		India	
	Total	Rural	Total	Rural
1973-74	66.2	67.3	54.9	56.4
1977-78	70.1	72.4	51.3	53.1
1983	65.3	67.5	44.5	45.7
1987-88	55.6	57.6	38.4	39.1
1993-94	48.6	49.7	36.0	37.3
2004-05	46.6	46.8	37.2	41.8
2009-10	37.0	39.2	29.8	33.8
2011-12	32.6	35.7	21.9	25.7

Source: Odisha Economic Survey-2014-15

In recent years a number of studies have been conducted around the world to identify the determinants of household poverty. Sikander and Ahmed (2008) used logistic regression analysis model for determining the factors responsible for the household level poverty in Punjab. Dudek and Lisicka (2013) applied the binary logit models approach to identify the households in danger of poverty. Achia *et al.*(2010), and Rusnak (2012) have shown that the logistic regression may be the appropriate choice, satisfying logicality and validity of certain assumptions inherent in the use of the model. In this study attempts have been made to find out significant variables affecting the rural poverty in Odisha. Mohanty (2016) has made a detailed discussion on the poverty scenario in Odisha along with small area estimation of the poverty at district level.

2. DATA

The data considered for poverty analysis of Odisha relates to the National Sample Survey (NSS) Central Sample household level primary data on consumer expenditure of the 68th round (2011-12),collected by the Government of India, National Sample Survey Office (NSSO). The Planning Commission, the Government of India has identified a poverty line based on expenditure data for each State for rural and urban areas separately. The poverty line has been defined as the minimum or the cut off standard of expenditure on food and non-food items or per capita income below which an individual or household is described as poor. The poverty line for rural Odisha as per the Tendulkar Methodology used for Odisha is Rs.695.00 monthly per capita expenditure (MPCE).

3. METHODOLOGY

In the 68th round of the National Sample Survey (NSS) on household expenditure, a stratified two stage sampling design was adopted. For sampling of the first stage units (FSU) Census villages and latest list of urban frame survey blocks formed by the NSSO were used for rural and urban areas respectively. For the urban sector, maximum 10 strata were formed within each NSS region. Each district of the State has been treated as a stratum. For each stratum a sample of FSU were selected in the form of two independent sub-samples by circular systematic sampling. FSU were sub divided into hamlet groups (hg) / sub blocks (sb) for rural / urban sector. For small FSUs whole village was listed for the selection of households. For large FSUs population more than 1200, two segments were formed on the basis of hg/sb. All the households listed in each segment were stratified in Second Stage Stratum (SSS). The sample households were selected from each selected FSU by using SRSWOR method. The total number of sample households for rural sector of the 68th round NSS was 2973. Poverty ratio explains the pattern of poverty, but is not concerned with explaining the causes of poverty.

Table 2. List of Explanatory Variables

Variable	Name of the Explanatory Variables	Variable type	Short form of Explanatory Variable
X_1	Household size	Discrete	hsize
X_2	Social group	Categories	sgrp
X_3	Total land possessed	Continuous	tot_pos
X_4	Sources of energy for lighting	Categories	light
X_5	Sources of energy for cooking	Categories	cook
X_6	Number of dwelling units	Discrete	dwell
X_7	Salary earner	Dichotomous	salary
X_8	Sex of the head of the household	Dichotomous	sex
X_9	Age of the head of the household	Continuous	age
X_{10}	Marital status of the household	Categories	m_stat
X_{11}	General education of the household	Categories	gen_edu
X_{12}	Percentage of MPCE under food items	Continuous	food_per
X_{13}	Total cultivated land	Continuous	land_cult

In this study a logistic regression analysis is made to identify the factors affecting poverty level of a household (whether a household is poor or not according to poverty line),which is considered as the dependant dichotomous variable by correlating with 13 independent (explanatory) y variables stated in Table 2.

3.1 Logistic Regression Model

Logistic regression model allows the prediction of a discrete outcome such as group membership from a set of explanatory (predictor) variables that may be continuous, discrete, categorical and dummy or a mixture of all these. In binary logistic regression model, the dependent variable is dichotomous like presence / absence of an attribute.

In present study the dichotomous dependent variable Y takes the value 1 (if the household is below poverty line) with probability p and takes the value 0 (if the household is above poverty line) with probability (1 – p).

The odds that Y=1 is defined as $\frac{P(Y=1)}{1-P(Y=1)} = \frac{p}{1-p}$, where p varies from 0 to 1 and the odds varies from 0 to positive infinity.

The natural logarithm of the odds $\log\left(\frac{p}{1-p}\right)$ is called the logit of Y, written as logit (Y) which varies from negative infinity to positive infinity. In logistic regression relationship between the dependent variable Y and the independent variable $X_1, X_2, X_3 \ldots\ldots, X_k$ is expressed as

$$\text{logit } (Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_k X_k$$

To convert logit (Y) back to the odds we have odds (Y=1)

$$= \frac{p}{1-p} = e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_k X_k}$$

On solving,

$$p = \frac{e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_k X_k}}{1+e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_k X_k}}$$

The logit function to be estimated is written as

$$\log \frac{p}{1-p} = \beta_0 + \sum_{i=0}^{k} \beta_i X_i$$

To estimate the (k+1) unknown parameters $\beta_0, \beta_1, \beta_2, \ldots\ldots, \beta_k$ from the observed data, the method of Maximum Likelihood is used because the dichotomous dependent variable makes estimation using Ordinary Least Squares (OLS) inappropriate. As Maximum Likelihood estimation process in case of logistic regression is not directly solvable, it is carried out by an iterative process.

$\log \frac{p_i}{1-p_i}$ is the natural log of the odds in favour of the i^{th} household falling below the poverty line where as β_i is the measure of change in the logarithm of the odds ratio of the chance of the poor to non-poor household and can also be written as :

$$\frac{\partial \log(\text{odds ratio})}{\partial X_i} = \beta_i$$

The logistic regression coefficient shows the change (increase when $\beta_i > 0$, decrease when $\beta_i < 0$) in the predicted logged odds of having the characteristic of intent for a unit change in the ith independent variable. In logistic regression the odds $(p/(1-p))$, which sometimes termed as odds ratio, represents the constant effect of a predictor X, on the likelihood that one outcome will occur. For instance, if the odds ratio is 1.15, there is an 15% increase in the odds at any value of X, and if the odds ratio is 0.95, there is an 5% decrease in the odds at any value of X. The odds ratio provides a single summary score of the effect.

The hypothesis testing about the coefficients is carried out with the help of Wald Test (instead of t-test), because the estimation method is not the standard OLS. This test is used to find out whether the explanatory variables are significant. The Wald statistic to test $H_0: \beta_i = 0$, is defined as $W_i^2 = [\hat{\beta}_i / \text{S.E}(\hat{\beta}_i)]^2$, which is asymptotically distributed as theoretical χ^2 with one degree of freedom.

For goodness-of-fit, -2 log likelihood (-2 LL), pseudo R^2 such as Cox and Snell R^2 and Nagelkerke R^2 may be computed.

Chi-square goodness-of-fit statistic is computed as $(-2LL_0)-(-2L\ L_m)$ which is distributed as χ^2 with m degrees of freedom, where m is the number of explanatory variables. If the chi-square is significantly contributes to the production of items the conclusion is at least one of the predictors.

As regards computation of R^2 found in ordinary least square method, the logistic regression does not have an equivalent one. There are many ways of calculating R^2 for the logistic regression, which are called 'Psuedo R^2'.

A Psuedo R^2 based on likelihood function may be computed as

$$R_L^2 = \frac{-2L\,L_0 - (-2L\,L_m)}{(-2L\,L_0)}$$

$$= \frac{L\,L_0 - L\,L_m}{L\,L_0}$$

The SPSS package gives two such Psuedo R^2 i.e Cox and Snell R^2 and Nagelkerke R^2.

Cox- Snell R^2 is defined as

$$Rcs = 1 - \left(\frac{L_0}{L_m}\right)^{2/n}$$

where L_0 is the value of the Likelihood function for a model with no predictors and L_m is the likelihood for the model being estimated with m predictors, and n is the sample size. A problem with Cox- Snell R^2 is that its upper bound is less than one. Hence, Nagelkerke R^2 is an adjusted version of the Cox- Snell R^2 that adjusts the scale of the statistics by dividing the Cox- Snell R^2 by its upper bound. 1- $L_0^{2/n}$. Besides, Cox- Snell and Nagelkerke Psuedo R^2 several other measures are available in logistic regression literature (see Long, 1997).

For fitting the logistic regression model, the computer software SPSS package has been used.

4. RESULTS AND DISCUSSIONS

4.1 Logistic Regression Model Fitting

The logistic regression model was estimated using the maximum-likelihood method of estimation using SPSS package and forward step-wise method to obtain the best model. This method automatically selects and includes the 'best' predictors into the model. In the analysis eight of the explanatory variables are found to be significant at 1% level of significance (Table 4). Three tests such as Chi-square test based on the -2 Log Likelihood, Cox and Snell R^2 and Nagelkerke R^2 are used to check the most significant variables (Table 3) indicated in the analysis.

Table 3. Model Summary

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
Null Model 3016.583	0.114	0.169
Fitted Model 2513.166	0.252	0.372

Source: Computed from primary data of NSSO

Table 4. Variables in the Equation

Variable	$\hat{\beta}$	S.E.	Wald Statistics	df	Significant (P-value)	Exp ($\hat{\beta}$)
hsize (X_1)	0.283	0.028	103.152	1	<0.001	1.328
sgrp(X_2)	-0.215	0.026	69.859	1	<0.001	0.807
light(X_4)	-0.313	0.027	136.215	1	<0.001	0.731
salary(X_7)	0.864	0.187	21.322	1	<0.001	2.374
age(X_9)	-0.016	0.004	12.357	1	<0.001	0.985
m_stat(X_{10})	0.466	0.141	10.977	1	<0.001	1.594
gen_edu(X_{11})	-0.140	0.019	54.347	1	<0.001	0.870
food_per(X_{12})	0.044	0.006	58.227	1	<0.001	1.045
Constant	-3.998	0.650	37.876	1	<0.001	0.018

Source: Computed from primary data of NSSO.

The results of the best model in the logistic regression analysis showing the estimated parameters with respective standard errors, Wald statistics, degrees of freedom, significant levels and Exp (Beta), the unique contribution of each predictor with corresponding P-values are presented in Table 4. The effects of the predictors are statistically significant at one percent level. Thus, the fitted logistic regression model is given by

$$\log\left(\frac{p}{1-p}\right) = -3.998 + 0.283 * \text{hsize} + (-0.215) * \text{sgrp} + (-0.313) * \text{light} + 0.864 * \text{salary} + (-0.016) * \text{age} + 0.466 * \text{m\_stat} + (-0.140) * \text{gen\_edu} + 0.044 * \text{food\_per}$$

The estimated slope coefficients (β) of the regression model shown in Table 4 to predict the log odds of the dependent variable but not directly the dependent variable as in OLS regression analysis. Each estimated slope coefficient gives the linear effect of a one-unit change in predictor variable on the log odds. By taking the antilog of both sides of the estimated model we have estimated odds as:

$$\frac{p}{1-p} = \exp[-3.998 + 0.283 * \text{hsize} + (-0.215) * \text{sgrp} + (-0.313) * \text{light} + 0.864 * \text{salary} + (-0.016) * \text{age} + 0.466 * \text{m\_stat} + (-0.140) * \text{gen\_edu} + 0.044 * \text{food\_per}]$$

The computed odds of predictor variables are shown in the column of Exp ($\hat{\beta}$) in Table 5. For example, the odds of variable household size (hsize) is 1.328. It means that the odds of a household being in the poverty are increased by 1.328 for each one unit increase in household size of the rural households. The probabilities (p) that Y = 1 is estimated as:

$$p = \frac{\exp[-3.998 + 0.283 * \text{hsize} + (-0.215) * \text{sgrp} + (-0.313) * \text{light} + 0.864 * \text{salary} + (-0.016) * \text{age} + 0.466 * \text{m\_stat} + (-0.140) * \text{gen\_edu} + 0.044 * \text{food\_per}]}{1 + \exp[-3.998 + 0.283 * \text{hsize} + (-0.215) * \text{sgrp} + (-0.313) * \text{light} + 0.864 * \text{salary} + (-0.016) * \text{age} + 0.466 * \text{m\_stat} + (-0.140) * \text{gen\_edu} + 0.044 * \text{food\_per}]}$$

This formula for p measures the probability that the i^{th} household being in poverty as a result of the different achievement levels of the given predictor variable. The computed odds and corresponding probabilities of each predictor variable (capability) are given in Table 5.

Table 5. Odds and Probability

Variables	$\hat{\beta}$	Odds of each Capability [Exp ($\hat{\beta}$)]	1+odds	Probability	
				p	1-p
hsize (X_1)	0.283	1.328	2.328	0.5704	0.4296
sgrp(X_2)	-0.215	0.807	1.807	0.4466	0.5534
light(X_4)	-0.313	0.731	1.731	0.4223	0.5777
salary(X_7)	0.864	2.374	3.374	0.7036	0.2964
age(X_9)	-0.016	0.985	1.985	0.4962	0.5038
m_stat(X_{10})	0.466	1.594	2.594	0.6145	0.3855
gen_edu(X_{11})	-0.140	0.870	1.870	0.4652	0.5348
food_per(X_{12})	0.044	1.045	2.045	0.5110	0.4890
Constant	-3.998	0.018	1.018	0.0177	0.9823

Analysis of odds

India, like the other developing countries is subject to the threat of high population growth rate. This high growth accompanied by high unemployment rate possesses a serious threat to wellbeing of the household. The household size contributes to high probability of becoming a poor household. The computed odds shows that it is 32.8% more likely that the household size affects the poverty level. Also if the family has a salary earner, this will help towards the reduction of household poverty. It is 137% more likely that the salary earner of the household affects the poverty level. It is often seen that if the head of the household is educated the descendants will also be likely to get reasonable education. It is 13 % less likely that the general education level of the head of the household affects the poverty level. Also, it is 27% and 19% less likely that the sources of energy for lighting and social group of the head of the household respectively affect the poverty level of that household in rural Odisha.

Marital status (m-stat) of the head of the household is an important factor for reducing the probability of remaining as poor household. It is observed from the analysis that it is 60% more likely that the marital status of the head affects the poverty level. Further, it is less than 2% likely the age of the head of the household affects the poverty level and it is 4.5% more likely that the per capita food expenditure affects the poverty level.

5. SUMMARY AND CONCLUSIONS

The determinants of poverty in Odisha have been explored through logistic regression analysis taking poverty (poor or non-poor) as a dichotomous variable. It is observed that household size, the level of education, age and marital status of the head of the households, salary earners in the household, sources of energy for lighting and social group of the household are the significant factors affecting the rural poverty in Odisha. It is surprising to note that the factors like total land possessed by the household and total cultivated land of the household have no significant effects on the rural poverty in Odisha, but have only marginal effect, although these two factors are the key variables which influence the standard of living in rural Odisha. It may be reasonably commented here that perhaps the selected sample households are not a representative sample for the analysis.

Keeping in view the factors affecting poverty in rural Odisha, strategic plans and programmes should be formulated by the Government for successful implementation of the poverty reduction programmes.

ACKNOWLEDGEMENTS

The authors immensely thank the esteemed referee for his valuable comments which improved the presentation of the paper to a great extent.

REFERENCES

Achia, T. Wangombe, A. and Khjadioli, N. (2010): "A Logistic Regression Model tro Identify Key Determinants of Poverty using Demographic and Health Survey data", *Europ. J. Soc. Sci.*, **13(i)**: 38-45.

Dudek, H. and Lisika, I.(2013): "Determinants of Poverty-Binary Logit Model with Interaction terms Approach", *Econometrics* **3 (41)**, PP 65-77.

Mohanty, B.N. (2016): "Regional Disparities in Orissa – A Multivariate Statistical Study," thesis submitted for the Ph.D. degree in Statistics, Utkal University, Bhubaneswar, Odisha.

Long, J.S. (1997) : "Regression Models for categorical and limited dependent variables analyis and interpretations", Thousands Oaks, CA : Sage Publication.

Rusnak, Z.(2012) : "Logistic Regression Models in Poverty Analysis", Econometrics, Vol. 35 (1), Publishing House of Wroclaw University of Economics, Wroclaw.

Sikandar, M.U and Ahmed, M. (2008) : "Household Determinants on Poverty in Punjab : A Logistic Regression Analysis of MICS (2003-04) Data set", 87th Global Conference on Business & Economics, Florence.