EUROPEAN ACADEMIC RESEARCH
Vol. VI, Issue 2/ May 2018

ISSN 2286-4822
www.euacademic.org

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

Estimation and Analysis the Impact of Demography Variables on Poverty in Sudan: A Logistic Regression Approach

ISMAIL ABAKER ADAM BANDH[1]
Faculty of Economic and Social Studies
University of Geneina, Western Darfur
HAMZA IBRAHIM HAMZA
Main supervisor, Faculty of Urban Sciences
University of Al-Zaiem Al Azhari
ALTAIYB OMER AHMED
Co-supervisor, College of Science
Sudan University of Sciences and Technology

Abstract:

Poverty has exposed out to be a great global economic and social problem; in Sudan, it is many-sided and deep-rooted. This study attempts to analyze the influence of demographic factors of households on poverty in Sudan, using the lasted National Household Budget and Poverty Survey data (2015). The study utilizes a logistic regression model was estimated to determine which variables might be important in explaining poverty in Sudan. The dependent variable is the probability of a household being poor or not and a set of demographic variables as the explanatory variables. Different households are classified as either poor or non-poor on the basis of a threshold yearly per capita spending of SDG 6082 and a daily calorie intake of 2110 calories. The poverty reveals that nearly 36.1% of the sample households live below poverty line and poorest state about 67.2%. The results showed that household size, sex and age of household head and dependency ratio significantly explains the poverty, whereas place of residence is not significant in explaining the household poverty status.

[1] Corresponding author: bandh77@gmail.com

We found that household size and dependency ratio were positively associated with the probability of being poor, and sex, an age of household heads were negatively related to the poverty status. Moreover, the value of statistics Nagerlelke R^2 explains 27.8% of the variance in the outcome. Results from this study revealed that in future, the poverty may be predicted by considering these identified influential variables.

Key words: Poverty, Demographic, Variables, Household, Economic, Logistic Regression.

1. Introduction

Poverty is considered as a sociological, economic, political and historical phenomenon facing individuals and families in different societies. In fact, poverty is the most risk, complicated and most measured issues, and despite its widespread and increasing in all of the developing countries, in Sudan, it became a terrifying matter. According to the World Bank, in the next 25 years, the world's population will roughly increase by additional 2 billion. About 97% of this increase will be in developing countries World Bank, (2000). The increasing population will be the problems of unemployment and poverty. The diminishing of poverty is currently a key issue of Sudan and all developing countries.

2. Poverty in Sudan

Poverty has been identified as a key challenge of human development particularly developing countries; whereas attempts have been made to understand and tackle it poverty prevalence has continued to increase over the years. Monitoring the poverty situation in Sudan has exposed that "poverty in Sudan as measured by all poverty ratios has increased rather significantly between 1978 and 1986. In an absolute sense, the consequences confirm that the number of poor households increased from 1.6 to 2.6 million respectively during that years, thus recording yearly rate of increase of 6.2%" Central Bureau of Statistic (CBS), (1992). According to (CBS), 2009 the National Household Budget and Poverty Survey (NHBPS) 46.5% of households in Sudan live below the poverty line. This represents about 14.4 million people and the last survey in 2015 may convey poverty line 36.1%. The poverty line computed from the Sudan Integrated National Household Budget and Poverty Survey data using the cost of basic needs method. We find that the poverty line declined in 2015 because the previous year's global standards were applied in measuring poverty but in the present study local calorie 2110 were used to measure poverty in Sudan.

According to CBS, (2015) in Sudan the present NHBPS2015 is the fourth in a series of the similar type of surveys undertaken by CBS; earlier conducted in 1967, 1978 and 2009. The current survey is, however, the first to be implemented after the secession in 2011.

The NHBPS2015 is designed to provide information over a wide range of topics feeding into socioeconomic statistics for all 18 states of Sudan.

Generally, poverty indicates a status of absolute lack of one or more of dimensions of welfare for an individual, such as the lack of access healthcare services, education, decreased human capital, the insufficiency of accommodation infrastructure, malnutrition, and lacking some of merchandise and services, the inability of expressing political opinions, and the like, and each one of these dimensions deserves a separated consideration. Also, poverty has nonmonetary dimensions, accompanied not only by a lack of the income or expenditure but also decline in social relations and insecurity. Mustafa., (2011).

Poverty is affected by many variables which are different from one community to another, that the researcher sensed the problem from the subsistence reality of many families and individuals whom their patterns of living and demographic characteristics have changed. Because of that, poverty determinants studies are important to providing a solution to this challenge in Sudan, which is aimed to achieve social and economic development more realistic. The basic idea behind this paper is to bring to light the various factors hindering the success of poverty alleviation in Sudan. The going up rate of poverty with the passageway of time is to be examined in relation to demography factors to determine the phenomenon association statistically.

In this study, we analyze the variables affecting poverty by considering the information we obtained from the Central Bureau of Statistic (CBS) of National Household Budget and Poverty Survey NHBPS., (2015) in Sudan. By considering the binary poverty household status of the response variable, the Logistic Regression approach is used to identify the variables that significantly affect the poverty. The rest of the paper is organized as follows. In Section 3, we review previous studies, in section 4, research objectives, in section 5, describe the statistical methods used in this paper.

Results and discussion are given in Section 6. Finally, we provide some conclusion and recommendation in Section 7.

3. Previous Studies

Current literature suggests numerous ways of modeling the determinants of poverty. Thus, there is no consensus for the selection of a model. The best analyzes to identify the factors that influence the probability of being poor is the regression analyses when we can check at the same time the influence of the different factors.

Habyarimana, Zewotir, and Ramroop, (2015). Analysis of demographic and health survey to measure poverty of household in Rwanda. We used the principal component analysis (PCA) technique in order to create the asset index. Then the asset index was used to assess the socio-economic status (SES) of households. The methodology is applied and demonstrated using the household survey data in Rwanda. The Rwanda data analysis showed that the age, gender of the household head, education level of the household head, and place of residence, the province of household head and size of the household were the significant predictors of poverty of the family in Rwanda.

Yusuf, (2015). Study about, Determinants of Rural Poverty in Tanzania this study target at assessing the determinants of poverty in Mkinga district in rural Tanzania. The ordinal regression model was used to model events of observing scores of livelihood status in the area of study. The study exposed that just about 93% of respondents in the area were poor. Gender, size of land the household owns, the size of the farm used in farming, Household size and the dependency ratio were found to be related to poverty, therefore influencing poverty in the area of study.

Deressa and Sharma, (2014). Study about, Determinants of Poverty in Ethiopia. The study attempts to analyze the impact of socioeconomic and demographic characteristics of households on poverty in Ethiopia, using the Household Income, Consumption and Expenditure Survey (HICES) 2010. The study uses a logistic regression model to identify determinants of the well-being of the household by considering per capita consumption as a dependent variable. The results reveal that female-headed households, large family size, and high dependency ratio are adversely affected.

The above previous review is important in understanding the findings over time of previous poverty studies that have been done. The review reveals that the logistic regression model approach is a more popular approach in poverty studies. The approach has a number of positive features and significant predictors of poverty in comparison to the expenditure approaches in studying poverty. The income is generally the measure of preference in developed countries. In the developing countries, expenditure is choice over income due to the difficulties involved in the measuring income.

Most of the previous studies focused on the measure the poverty by income specific in Sudan this includes Abaker, and Salih, (2012). This study seeks to bridge this gap.

Household size and sex of head and place of residence were identified by several studies as an important factor affecting poverty. This study contributes to poverty in Sudan by identifying the unique determinants such as the age of head and dependence ratio as an explanatory variable, which other previous studies did not focus on in Sudan.

4. The Objectives of the Study:

1. To identify and analysis the determinants that affect household poverty status.
2. To estimate and examine the impacts of demography factors growth on poverty.

5. The Methodology of the Study

This research analyses the NHBPS data using the SPSS statistical Packages (version 2o). The logistic regression model main considers in the build and analysis the model of the study. The level of significance used in all the statistical tests run is the conventional 5%.

5.1. Data Sources

This paper depends totally on qualitative and quantitative raw data that collected from National Household Budget and Poverty Survey (NHBPS) undertook by Central Bureau of Statistics - Sudan (2015), which is the most recent, obtainable at the time this study is written.

A sample of 690 clusters was selected at the third phase for each of the 18 states of Sudan, with the total of 13800 households for Sudan, the available data of the researchers to use only 5965 households.

5.2 Variables of study

For logistic regression model used in this study, household expenditures per capita were measured as average household income in SDG per year are considered dependent variables. This is calculated considering both food and non-food expenditure including in-kind values in the household. These were codified in the poor (1) and non-poor (0). Families with yearly per capita consumption expenditure less than the poverty line are considered poor and those with costs greater than the poverty threshold are considered non-poor.

The set of independent variables, that are included in the model of the determinants of poverty in Sudan some important demography factors include place of residence, household size, sex of household head, the age of the Household head and dependency ratio.

5.3 Logistic regression model

Logistic regression is a method of modeling the dependence of a binary response variable which takes values 1 and 0. Logistic regression gives each predictor a coefficient which measures its independent contribution to variation in the dependent variable.

Model Assumptions: Firstly, explanatory variables do not need to be normally distributed. It works better where the group sizes are very disparate. Secondly, it does not assume a linear relationship between the dependent and independent variable, but a linear association between the logit of the response and explanatory variables. Thirdly, the error term is independent and there is no assumption of a normal distribution. Finally, independent variables don't have strong co-linearity. Mathematically the resulting models are easier to interpret due to its mathematical simplicity Mbugua, (2014).

The dependent variable Y takes the value 1 if the household is 'poor' and takes a value 0 if 'non poor'.

Let P=P(Y=1) then 1-P=P(Y=0) and the logistic regression model is defined as follow

$$\log\left(\frac{p}{1-p}\right)= B_0+B_1X_i+B_2X_{2i}+\ldots+B_kX_{ki} \quad \ldots\ldots\ldots\ldots \quad (1)$$

Where β_0 is the intercept and β_k is the regression coefficient of the i^{th} predictor, i= 1, 2,..,k.

The odds ratio is a measure of effect size, describing the strength of the relationship or non-independence between two binary data values. It treats the two variables being compared symmetrically and can be estimated using some type of non-random samples. It is used as a descriptive statistic and plays an important role in logistic regression. Odds ratio another way out to interpret the coefficients is by considering only the ratio of the probability that the event occurs and the probability that the event does not occur Gerolimetto, (n.a)).

This ratio is:

$$\frac{p(Y=1)}{p(Y=0)} = \frac{\frac{e^{XB}}{1+e^{XB}}}{\frac{1}{1+e^{XB}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

In this case, the exponentiated coefficients echo changes in the odds ratio, consequently to a unit difference in the explicative variable. Coefficients (β) are in particular helpful to determine the sign of the relationship: a positive coefficient indicates that a unit increase in the X is connected with increases the predicted probability and vice versa.

With logistic regression we model the natural log odds as a linear function of the explanatory variables:

Logit (y) = ln(odds)= $\ln(\frac{p}{1-p}) = B_0 + B_1X_1 + B_2X_2 + \ldots + B_kX_k$. (3)

Where p is the probability of interested outcome and x_k is the explanatory variables. The parameters of the logistic regression are B_0 and B_k. Let P_i denotes the probability that the i[th] household is below the poverty line. We assume that P_i is a Bernoulli variable and its distribution depends on the vector of predictors X. If probabilities of the event of interest happening for individuals are needed, the logistic regression equation can be written as follow:

$$P = \frac{\exp(\beta_0+\beta_1X_1+\beta_2X_2+\cdots\ldots+\beta_kX_k)}{1+\exp(\beta_0+\beta_1X_1+\beta_2X_2+\cdots\ldots+\beta_kX_k)} \quad , o< p < 1 \;\ldots\ldots\ldots \quad \ldots\ldots\ldots \quad (4)$$

Where, p = the probability that a case is in a particular category, exp = the base of natural logarithms (approx 2.72), B_0 = the constant of the equation and, B_k = the coefficient of the predictor variables.

In the paper are considered a set of variables that could influence the probability of being poor.

The binary logistic regression shows how these independent variables affect the increase of poverty. So as a dependent variable is considered variable poor taken value "1" if are poor and "0" if it is not poor. As independent variable are considered: Place of residence (rural =1, urban =0), household size, sex of household head (female=1, male=0), the age of household head and dependency ratio.

Following is the multiple regression model specification:

$$P_i = B_0 + \beta_1X_i + \beta_2X_{2i} + \ldots\ldots + \beta_kX_{ki} + \varepsilon_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

Where Pi denotes probability, B_0 the constant term, β k parameters that will estimate, xi is a vector of independent (explanatory) variables, ε error term.

$$\log\left(\frac{p}{1-p}\right)=B_0+B_1PR+B_2HS+B_3SHH+B_4AHH+B_5DR+□_6\ldots\ldots\ldots\ldots\ldots(6)$$

According to Mbugua, (2014). Automatic stepwise selection procedure:

Models building strategy of binary logistic regression analysis has **entered** a procedure for variable selection in which all variables in a block are entered in a single step.

5.3.1 Maximum Likelihood Estimation

The maximum likelihood estimate is that value of the parameter that makes the observed data most likely. The logistic regression model just developed is a generalized linear model with binomial errors and link logit. We can, therefore, rely on the general theory developed in logistic regression to obtain estimates of the parameters and to test hypotheses.

Small (Finite) sample tests, e.g., t-test and F-test, cannot be used to test hypotheses in the linear probability model. This is because the error term has a binomial distribution, not a normal distribution. To test hypotheses, you must use large sample (asymptotic) tests. These include the t-test, approximate F-test, Likelihood ratio test, Wald test, and Hosmer-Lemeshow Test. MLE allows more flexibility in the data and analysis because it has fewer restrictions Hosmer and Sturdivant, (2013).

We want to choose β's that maximizes the probability of observing the data we have:

$$L=\Pr(y_1,y_2,\ldots,y_N)=\Pr(y_1)\Pr(y_2)\ldots\Pr(y_N)=\prod_{i=1}^{N}\Pr(y_i) \quad \ldots\ldots\ \ldots\ldots\ldots\ldots (7)$$

Substituting in using logistic regression model:

$$\ln L=\sum_i y_i\beta x_i-\sum_i \ln(1+\exp(\beta x_i))$$

5.3.2 Pseudo R^2 Measures for Logistic Regression

Most statistical packages provide further statistics that are used to measure the usefulness of the model and that are similar to the coefficient of determination (R^2) in linear regression. The Cox and Snell and the Nagelkerke R^2 are two such statistics.

5.3.3 Cox-Snell R^2

In linear regression using ordinary least squares, a measure of goodness of fit is R^2, which represents the proportion of variance explained by the model. Using logistic regression, an equivalent statistic does not exist, and therefore several pseudo-R^2 statistics have been developed. The Cox-Snell R^2 is a pseudo – R^2 statistic and the ratio of the likelihoods reflects the improvement of the full model over the intercept-only model with a smaller ratio reflecting greater improvement Hosmer, Lemeshow, and Sturdivant, (2013).

It is given by:

$$\text{Cox-snell } R^2 = 1-\left[\frac{L(R)}{L(F)}\right]^{2/N} \qquad \text{...... ....(8)}$$

Where, L(R) = likelihood of the intercept-only model, L(F) = likelihood of the specified model, N = Number of observations.

5.3.4 Nagelkerke R^2

The *Nagelkerke* R^2 adjusts the Cox-Snell R^2 so the range of possible values extends to one Hosmer, Lemeshow, and Sturdivant, (2013).

$$\text{Nagelkerke } R^2 = \frac{1-\left[\frac{L(R)}{L(F)}\right]^{2/N}}{1-L(R)^{2/N}} \qquad \text{.. (9)}$$

Where, L(R) = likelihood of the intercept-only model, L(F) = likelihood of the specified model, N = Number of observations.

5.3.5 Wald Test

According to Mbugua, (2014) Wald χ^2 statistic is used to test the significance of the individual regression coefficients in the model. It is calculated as:

$$\text{Wald} = [\frac{\hat{B}}{\hat{\sigma}_{\hat{B}_i}}]^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(10)$$

where $\hat{\sigma}_{\hat{B}_i}$ is an estimate of the standard error of b provided by the square root of the corresponding diagonal element of the covariance matrix, V($\hat{B}$). Each Wald χ^2 statistic is compared to a Chi-square distribution with 1 degree of freedom. This method of reliability is questionable, particularly for small samples. Likelihood ratio tests are considered superior.

5.3.6 Likelihood Ratio Test

This test for a particular parameter compares the likelihood of obtaining the data when the parameter is zero (L_0) with the likelihood (L_1) of obtaining the data evaluated at the maximum likelihood estimate of the parameter. The test statistic is calculated as

$$LR = -2[L(F)-L(R)] \quad \ldots\ldots\ldots \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \qquad (11)$$

Where, L(F) the maximized log-likelihood of the full model and L(R) the maximized log-likelihood of the reduced model

Full Model

$$\text{Logit } [\pi(x)]= \ln\left(\frac{\pi(x)}{1-\pi(x)}\right)= B_0+B_1X_1+\ldots+B_{q-2}X_{q-2}+B_{q-}X_{q1}+B_qX_q+B_{q+1}+\ldots+B_{p-1}$$

Reduced Model

$$\text{Logit } [\pi(x)]=\ln\left(\frac{\pi(x)}{1-\pi(x)}\right)= B_0+B_1X_1+\ldots+B_{q-2}X_{q-2}+B_{q-1}X_{q-1} \quad \ldots\ldots..(12)$$

The null and alternative hypotheses with respect to this test are shown below.

H_0: $B_q= B_{q+1} = \ldots= B_{p-1}= 0$

H_a: not H_0

This LR statistic is asymptotically distributed as χ^2 with *p*-*q* degrees of freedom. Or the Likelihood Ratio Statistic:

$$G^2 = -2\log_e\left[\frac{L(R)}{L(F)}\right] = -2[\log_e L(R) - \log_e L(F)]\ldots\ldots\ldots\ldots\ldots(13)$$

The Decision rule: If $G^2 \leq \chi^2(1\text{-}\alpha;p\text{-}q)$, conclude H_0, If $G^2 > \chi^2(1\text{-}\alpha;p\text{-}q)$, conclude H_a. It is compared with a chi-square distribution with 1 degree of freedom. It indicates whether the parameter contributes significantly in predicting the dependent variable.

5.3.7 Goodness of Fit of the Model- Hosmer - Lemeshow test

This measures how well the model describes the response variable. Assessing goodness of fit involves investigating how the values predicted by the model are close to the observed values.

Hosmer-Lemeshow test is commonly used for assessing goodness of fit of a model and allow for any number of explanatory variables which may be continuous or categorical. The observations are partitioned into groups of approximately equal sizes. The observations are grouped into deciles based on predicted probabilities Mbugua, (2014). The test statistic is calculated using the observed and expected counts for the categories as:

$$H = \sum_{g=1}^{10} \frac{(Og-Eg)^2}{Eg} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (14)$$

Where *Og* and *Eg* denote the observed events and expected events for the g^{th} risk deciles group. Small values indicate a good fit to the data, therefore, good overall model fit.

Large values (with $p < .05$) indicate a poor fit to the data. Hosmer and Lemeshow do not recommend the use of this test when there is a small *n* less than 400 Hosmer & Lemeshow, (2000).

6. RESULTS AND DISCUSSIONS

The logistic regression techniques have been applied to evaluate the demography characteristics of the household's head and household characteristics as the determinants of household poverty in Sudan. The definition of the demography characteristics and outputs are as shown in the tables (1, 2, 3, 4 and 5).

Definition of the dependent variable we classified the household as either poor or non-poor dazed on their per capita expenditure. Namely, two categories are represented by one binary variable, takes the value 1 if the household is poor and o if it is not poor according to poverty line that is a household is considered to be poor if its total consumption is below the poverty line.

Table (1) definitions of the demography variables

N	Explanatory Variable	Abbreviations	code	Definition	Characteristic
1	Place of residence	PR	X_1	whether a household is located in the rural or urban area	2 categories are represented by one binary variable as follows: Dummy, Rural = 1, Urban = 0
2	household size	HS	X_2	Total household members	Continuous
3	sex of household head	SHH	X_3	Sex of household head (male or female).	Dummy, Female = 1 , Male =0
4	age of household head	AHH	X_4	Age of household head (years)	Continuous
5	dependence ratio of household	DR	X_5	DR of number of members (<15 years and >64 years) to household size and treated as continuous variable.	Continuous

According to descriptive analysis, the average of household size and dependency ratio were found to be 7.37 and 163.05 percent respectively. The other demographic factors gender and age the results shows that 31.1% of the more poorest households are aged between 35 and 44 as compared to 25.47% of the poor households head who are over 54 years and 16.28% of the poor who are less than 35 years, and more poor household head in age 40, while the average of age was found to be 46.04 of the poor.

Table (2) Crosstabs of Pearson chi-square statistics test and Likelihood Ratio for the association between demographic characteristics with Poverty

Variable	Pearson Chi-Square	Likelihood Ratio	df	Sig
PR	6.352	6.401	1	.012
HS	1082.681	1171.799	17	.000
SHH	1.041	1.o35	1	.311
AHH	97.531	98.634	8	.000
DR	515.114	527.017	10	.000

Source: Prepared by researcher from the Survey Data, 2015

To test the association of the variables, in this section we apply the Chi-square test. To perform this, we compare all the explanatory variables with response variable, poverty. The results of the tests are shown in the table (2), we observe that there is a very strong association between places of residence, household size, age of household head, and dependency ratio with poverty household status, except for the variable sex of household head was not significant at 5% level of significance, but it's significant when included all variables in the logistic regression. Also, the place of residence is not significant when included in the logistic regression.

Table (3) Coefficients and Wald tests for logistic regression on the poverty and demography data

	B	S.E	Wald	df	Sig	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step1[a] PR(1)	-.065	.070	.864	1	.353	.937	.817	1.075
HS	.410	.015	718.322	1	.000	1.507	1.463	1.553
SHH(1)	-.566	.093	37.402	1	.000	.568	.473	.681
AHH	-.049	.024	4.303	1	.038	.952	.909	.997
DR	.160	.015	113.782	1	.000	1.173	1.139	1.208
Constant	-3.162	.150	444.194	1	.000	.042		

Source: Prepared by researcher from the Survey Data, 2015

Variable (s) entered on step 1: PR, HS, SHH, AHH, and DR. a

The purpose of this paper was to identify the demography factors affecting poverty by using binary logistic regression model in Sudan.

This method was adopted in line with other studies by Sikander, and Ahmed, (2008). Garza-Rodriguez, Zyka, and Bici, (2015) and (2014)., Ngunyi, (2015).

According to the summary statistics from the table (3) shows that out of the fifth identified variables only one variable was not significant in explaining whether household's status is poor or not poor. Also, we see that household size (HS) is a statistically significant variable (B= 0.410, Wald = 718.322, P =.000) and its positive coefficient indicates that with increasing household size increases the probability that the household be poor. This finding is consistent with previous studies in Habyarimana, Zewotir, and Ramroop, (2015) And Garza-Rodriguez, (2015) who found that having more dependent members and more household members in general, seems to reduce per capita income. Thus, even though the coefficient (B= -.566, Wald=37.402, P-value=.000) for the Sex of household head (SHH) variable is negative and statistical significance. The gender of the head variable is an important factor in explaining the poverty status of the family but the negative coefficient indicates that households headed by a female have the lower probability of being poor than male-headed households.

The coefficient for the variable age of household head (AHH) is negative and statistically significant variable (B= -.o49, Wald= 4.303. P=.000). This means that there is an established negative relationship between age of household head and the per capita expenditure of the household. The age of household heads grows older; the per-capita expenditure/income of the household reduces, thus, increasing the level of poverty in the household. Thus, we can see that an increase of one year in the age of the head decreases the odds of being poor by almost 95.2%. We found that there is a strong and statistically significant inverse relationship between poverty and age of the head.

The coefficient of the dependent ratio (DR) is positive and statistically significant (B=.160, Wald= 113.782, P= .000). The odds ratio of the variable dependency ratio shows a contribution of 20.8% in increasing the likelihood of being poor whereas household size (HS) contributes 55.3%.

Therefore a majority of households fell into poverty because of having large families with many dependants being children or elderly at unproductive age.

Type of place of residence (PR) had an insignificant impact on poverty. It had been found that (B = -.o65, Wald = .864, p = .253). This means that there is a negative relationship between the type of place of residence and poverty. Theoretically, we can also explain this fact, as most of the household are residence in the urban areas and rural areas influenced by the poverty according to our income and levels of per capita expenditures in Sudan.

While, as can be seen in the table (3), the variables HS and DR have odds ratios greater than one, which means that these variables are positively correlated with the probability of being poor, while those variables SHH and AHH which have odds ratios lower than one are inversely correlated with the probability of being poor. The confidence interval for the odds ratio of PR includes the number one, which means that this variable has no statistically significant effect on the probability of poverty.

Table (4) Omnibus Tests of Model Coefficients

	chi-square	Df	sig
Step 1 Step	1305.466	5	.000
Block	1305.466	5	.000
Model	1305.466	5	.000

Source: Prepared by researcher from the Survey Data, 2015

In the above table (4) we have added all five explanatory variables in one block and therefore have only one step. This means that the chi-square values are the same for step, block, and model. Here the chi-square is highly significant (chi-square=1305.466, df = 5, p-value <.000) so our model is significantly better, which indicates the accuracy of the model improves when we add our explanatory variables.

Table (5): Summary measures of goodness-of-fit statistics an of the model with selected covariates

Summary Statistic test	**Value**	**df**	**P-value**
Hosmer - Lemeshow	32.654	8	.000
LR chi-square	1305.47	8	.000
Log-likelihood	3044.101		
Cox and Snell R^2	0.199		
Nagelkerke R^2	o.278		
Pseudo R^2	0.1766		

Source: Prepared by researcher from the Survey Data, 2015

According to the table (5), the variables are significant predictors of poverty ($p< \mathbf{0.05}$) Goodness-of-fit statistics assess the fit of a model against actual values. The inferential goodness-of-fit test is the Hosmer-Lemeshow (H-L) test that yields a $\chi^2$of 32.654 and was significant Suggesting that there was lack of fit the model of the data. Thus we not accept the null hypothesis that household characteristics and perceptions have not influence on poverty. The log likelihood yields a $\chi^2$of 3044.101 and was significant at (p< $\mathbf{0.05}$) which also give a good fit for the model to the data and thus the null hypothesis was also tenable for the model. Values of statistics Cox-Snell R^2 and Nagerlelke R^2 parameters are 0.199 and 0.278 which indicate that the model explains 19.9% to 27.8% of the variance in the outcome. This low value is explained primarily by the fact that the main variable affecting the poor are household income, these variables.

The model has a pseudo R^2 of 0.177 which means that 17.7% of the variation in the dependent variable is due to the variations in the independent variables.

Table (6) Correct Classification Table [a] of the Model

Observed		Predicted		
		Poverty household status		Percentage Correct
		Non-Poor	poor	
Poverty household status	Non-Poor	3541	449	88.7
	Poor	1101	793	41.9
Overall Percentage				73.7

Source: Prepared by researcher from the Survey Data, 2015
The cut value is .500[a]

The classification table (6) shows that the model makes a correct prediction 73.7% of the cases compared to 67.8% in the null model, a marked improvement. Of the 3990 households with non-poor, the model correctly identified 3541 of them as not likely to have one. Similarly, of 1894 that did have a poor, the model correctly identifies 793 as likely to have one. 41.9% is also known as the **sensitivity** of prediction. 88.7% is also known as the **specificity** of prediction.

The logistic model was fitted to the data to test the relationship between the likelihood of a household being poor or non-poor. The logistic regression analysis was carried out by entering method, and the result showed that.

The optimal model:

$$\text{Log}\left(\frac{P}{1-P}\right) = Y = -3.162 - 0.65\ PR\ (1) + 0.410\ HS - 0.566\ SHH\ (1) - 0.049\ AHH + 0.160\ DR \quad \ldots \ldots (15)$$

$$P = \frac{\exp(Y = -3.162 - 0.65\ PR\ (1) + 0.410\ HS - 0.566\ SHH\ (1) - 0.049\ AHH \quad + 0.160\ DR)}{1+\exp(Y = -3.162 - 0.65\ PR\ (1) + 0.410\ HS - 0.566\ SHH\ (1) - 0.049\ AHH \quad + 0.160\ DR)},\quad 0 < p < 1 \quad \ldots \ldots (16)$$

The estimates of the logistic regression are shown in the above Tables. In general, the logit model fitted the data quite well. The chi-square test strongly rejects the hypothesis of no explanatory power and the model correctly predicted 73.7% of the observations.

Furthermore, household size, sex of household head, the age of household head and dependency ratio are statistically significant and the sign on the parameter estimate support expectations, while the variable place of residence is not significant.

7. Conclusions and Recommendations

7.1 Conclusions

Therefore we can conclude from the results reported above that:

- The result reveals that household size, sex of household head, the age of household head and dependency ratio significantly explains the poverty status of a household.
- Type of place of residence (PR) had the insignificant impact on poverty.
- The household size and dependency ratio were positively associated with the probability of being poor.
- The sex of household head and age of household head were negatively related to the poverty status.

7.2 Recommendations

The analysis undertaken in this study leads to the following guidelines implications for the researcher and government:

- ✓ Results from this study revealed that in future, the poverty may be predicted by considering these identified influential variables.
- ✓ We recommend using logistic regression to measure the impact of different poverty factors by utilizing per capita expenditures particularly in developing countries.
- ✓ We also recommend using the neural network models and comparing it with the logistic regression model to measure the impact of different poverty factors.
- ✓ This study recommends a careful review on the reforms to be taken in relation to household size and dependency ratio, suggests that expansion of education and intensification of family planning programmed at the grass root level are amongst rural areas.
- ✓ The government should be focusing on improving the livelihood situation and health services.
- ✓ The government should look at the labor conditions of females and reduce poverty; attention must be paid to the manufacturing sector and agriculture.

REFERENCES:

[1] Abaker, A.M. and Salih, A.A., 2012, April. Poverty and inequality in Services Sector of Sudan. IJPSS Volume 2, Issue 4.

[2] Achia, T. N. O., Wangombe, A. and Khadioli N. 2010. "A Logistic Model to Identify Key Determinants of Poverty Using Demographic and Health Survey Data", European Journal of Social Sciences 13, (1), 38-45.

[3] Ballon, P., and Duclos, J.Y., 2015. Multidimensional Poverty in Sudan and South Sudan.

[4] Central Bureau of Statistics; 2009. Poverty in northern Sudan, the estimates National Baseline Household Survey data, 2009, Khartoum, 2010.

[5] Deressa, T.K., and Sharma, M.K., 2014. The Determinant of Poverty in Ethiopia. Ethiopian Journal of Economics, 23(1), pp.113-130.

Developing Country Studies. China Agricultural University. No.17 Qing Hua Dong Lu. 5(6), 2015.

[6] Elhadary, Y.A.E., and Samat, N., 2016. Addressing Poverty in Sudan and Malaysia: A Story of Success and Constraints.

[7] Garza-Rodriguez, J., 2015. The determinants of poverty in the Mexican states of the US-Mexico border. Browser Download This Paper.

[8] Gerolimetto. M, Advanced data analysis Logistic Regression

[9] Habyarimana, F., Zewotir, T. and Ramroop, S., 2015. Analysis of demographic and health survey to measure poverty of household in Rwanda. Etude de la Population Africaine, 29(1), p.1472.

[10] Hosmer Jr, D.W., Lemeshow, S. and Sturdivant, R.X., 2013. Applied logistic regression (Vol. 398). John Wiley & Sons.

[11] Hosmer, D.W., and Lemeshow, S., 2000. Special topics. Applied Logistic Regression, Second Edition, pp.260-351.

[12] Mbugua, M.D., 2014. Application of Logistic Regression in Identifying key Determinants of Domestic Violence in Kenya (Doctoral dissertation, University of Nairobi).

[13] Mustafa, H. A; 2011. a book of "poverty in the context of globalization, University House, Alex 2011, PP 19.

[14] Ngunyi, A., 2015. Multidimensional Analysis of the Determinants of Poverty Indicators in the Lake Victoria Basin (Kenya).

[15] Puurbalanta, R. and Adebanji, A.O., 2016. Household Poverty-Risk Analysis and Prediction Using Bayesian Ordinal Probit Models. International Journal of Statistics and Applications, 6(6), pp.399-407.
[16] Sikander, M.U., and Ahmed, M., 2008. Household Determinants of Poverty in Punjab: A Logistic Regression Analysis of MICS (2003-04) data set. In 8th global Conference on Business and Economics.
[17] World Bank; 2000. World Development Report 2000/2001: Attacking Poverty, New York: Oxford University Press.
[18] Yusyf, H. M., 2015. The determinants of Rural Poverty in Tanzania.
[19] Zyka, E. and Bici, R., 2014. Identifying Household Level Determinants of Poverty in Albania Using Logistic Regression Model.
[20]http://documentation.statsoft.com/portals/0/formula%20guide/Logistic%20Regression%20Formula%20Guide.pdf.Statistica,Formula Guide: Logistic Regression2013.
[21] www.dst.unive.it/~margherita. 2018 .27