EUROPEAN ACADEMIC RESEARCH
Vol. VI, Issue 6/ September 2018

ISSN 2286-4822
www.euacademic.org

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

Estimation of the Social Variables Influencing Poverty by Using Logistic Regression Model to National Household Survey Data 2015 in Sudan

ISMAIL ABAKER ADAM BANDH[1]
HAMZA IBRAHIM HAMZA
Sudan University of Sciences and Technology
College of Science, Department of Statistics, Khartoum, Sudan

Abstract:

Poverty is a global phenomenon and has proven difficult to resolve and remained determinedly high in developing a country. The objective of this study is to determinate the most important social factors that influence poverty in Sudan. The data used in the paper come from the National Household Budget and Poverty Survey2015. A logistic regression model was estimated to determine which variables might be significant explaining poverty in this country. The dependent variable is the probability of a household being poor or not and a set of social variables as the explanatory variables. 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 result confirms that the variables negatively correlated and significantly explain the variations in the likelihood of being poor with the household poverty statuses are: highest level of school, marital status, and type of disability. While the read and write with understanding and vocational training are negatively related to the poverty status but not significant. This study recommends that government and households should be focusing at the significances variables found in this study

[1] Corresponding author: bandh77@gmail.com. University of Geneina, Faculty of Economic and Social Studies, Western Darfur.

and highlight the mechanisms that can be beneficial in reduction of the poverty levels in Sudan.

Key words: Poverty, Social, Variables, Household, Logistic Regression, Sudan.

1. INTRODUCTION

This paper investigates the determinants of poverty and household welfare in Sudan using the National Household Budget and Poverty Survey (NHBPS) data (2015). Poverty as a multidimensional concept includes monetary and nonmonetary characteristics. As a consequence of this multidimensional nature, there are used different measures which are different from one county to another. Even between people in a country observation for poverty change within regions and social and economic groups depending on their sources of income and determinants of well-being. People live in poverty when they are deprived of incomes and other living resources, such as goods, housing conditions, commodities and sanitation, services that can authorize them to have a role and build their own social life (Myftaraj, Zyka, and Bici, 2014).

In developing countries, poverty is very widespread and is characterized by hunger, lack of living resources, unemployment, illiteracy, epidemics, lack of health services and lack of water. In developed countries poverty is characterized by social exclusions, low salaries, and unemployment growth. Poverty measures are based on consumption and income (Myftaraj, Zyka, and Bici, 2014).

According to Castaneda, Doan, Newhouse, Nguyen, and Azevedo, (2016), in 1990, about 2 billion people, or 37% of the global population, lived on less than the current international poverty line of $1.90 a day. By 2013, the year for which the latest global poverty estimates are available, the number of extremely poor persons had fallen by over 60%. During the same period, the proportion of the global population living in extreme poverty fell even faster, from 37 to 11%. The Millennium Development Goal of halving extreme poverty in developing countries between 1990 and 2015 was met in 2010, five years ahead of time. Ending poverty in all its forms is the first of the 17 Sustainable Development Goals adopted by the United Nations, and the World Bank has set a determined goal of reducing the rate of extreme poverty to 3% by 2030.

2. POVERTY IN SUDAN

Seeing as Sudan's independence in 1956, the country has been wracked with unstable conflict. The obstinate aggression, an unforgiving climate and a tumultuous government controlled by military personnel are important causes of poverty in Sudan. Although Sudan's GDP per capita rose to $2,140 in 2016, unequal distribution of wealth and resources has exacerbated socioeconomic inequality through different regions of the country[1].

According to the Human Development Index of 2000, 26.6% of the population will likely not live past the age of 40. Nine out of 10 people live on less than a dollar per day. About 40% of the population lacks access to safe drinking water. Sudan has the largest number of internally displaced people in the world. Sudan ranks 147 out of 177 countries. Why is Sudan poor[2]?

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).

While Sudan still ranks among the poorest countries in the world, it has also seen great economic progress. Since 1990, Sudan's extreme poverty rate declined from 85% to 46% and continues to be on the decline until 2012, according to the results from the Millennium Development Goals. The reality is that one out of two people living in Sudan is impacted by poverty; meaning, he or she does not have the means to buy food or the ability to properly care for him or herself. Results from the MDGs expose that an estimated 15 million people all over Sudan are considered poor[3].

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.

But the results of the NHBPS survey 2015 show that the annual per capita consumption in Sudan was Sudanese Pounds (SDG) 6,082. Urban areas displayed average consumption levels higher than rural areas, at SDG 7,149 and SDG 5,509 respectively. Among States, average consumption was the highest in Khartoum, followed by Northern and River Nile. States of Darfur and Kordofan regions showed the lowest level. The NHBPS in 2015 may convey poverty line 36.1%. The poverty line computed from Sudan included National Household Budget and Poverty Survey data using the cost of basic needs method. The decline in the poverty line in Sudan 2015 is due to the change in calorie measurement method. In the past, international standards were used for calorie measurement, but in this study, local calories 2110 were used to measure poverty line, according to a briefing of the Central Bureau of Statistics. While, find that the highest poverty in Central Darfur and south Kordofan at 67.2% and 67% respectively, but the lowest in the northern state at 12.2%. To struggle poverty in Sudan, the World Bank agreed to provide $100 million in order to establish development projects in

Sudan until 2019. Sudan's state minister predicted that the economy of Sudan would grow by only 0.2% per year2. However, the situation remained unchanged in the economy, but now a day's worse than it was.

Generally, poverty affected by many factors which are different from community to another, that the researcher sensed the problem from the reality of many families and individuals whom their patterns of living and social 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 social factors to determine the phenomenon association statistically.

3. THE OBJECTIVES OF THE STUDY:

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

4. THE METHODOLOGY OF THE STUDY

This research analyses the NHBPS data using the SPSS statistical Packages (version 2o) and based on the descriptive, and analytically approaches. 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%.

4.1. Data Sources

This paper depends totally on qualitative and quantitative raw data that collected from National Household Budget and Poverty Survey (NHBPS) undertaken by Central Bureau of Statistics - Sudan (2015), which is the most recent, obtainable at the time this study is written. The total estimated population covered by the NHBPS 2015 is 34.2 million persons distributed on 6 million households. The survey does not include population groups such as nomads, people living in camps and homeless people etc and can thus not directly be compared to the population counted in the Census of 2008 or updates hereof. 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.

4.2 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). A household 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 social factors include can read and write with understanding, the highest level of school ever completed, ever attended vocational training, marital status and Suffering from any type of disability that prevents from doing usual work.

5. LOGISTIC REGRESSION MODEL

Logistic regression sometimes called the logistic model or logit model analyzes the relationship between multiple independent variables and a categorical dependent variable and estimates the probability of occurrence of an event by fitting data to a logistic curve. There are two models of logistic regression, binary logistic regression, and multinomial logistic regression. Binary logistic regression is typically used when the dependent variable is dichotomous and the independent variables are either continuous or categorical. When the dependent variable is not dichotomous and is comprised of more than two categories, a multinomial logistic regression can be employed (Park, 2013).

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 following

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

5.1. Odds

According to Park, (2013), the odds of an event are the ratio of the probability that an event will occur to the probability that it will not occur. If the probability of an event occurring is *p*, the probability of the event not occurring is (*1-p*). Then the corresponding odds is a value given by

$$\text{Odds of \{Event\}}=\frac{p}{1-p} \ldots\ldots (2)$$

Since logistic regression calculates the probability of an event occurring over the probability of an event not occurring, the impact of independent variables is usually explained in terms of odds. With logistic regression, the mean of the response variable *p* in terms of an explanatory variable *x* is modeled relating *p* and *x* through the equation $p=\alpha+\beta x$.

Unfortunately, this is not a good model because extreme values of *x* will give values of $\alpha+\beta x$ that does not fall between 0 and 1. The logistic regression solution to this problem is to transform the odds using the natural logarithm (Peng, Lee, and Ingersoll, 2002). With logistic regression we model the natural log odds as a linear function of the explanatory variable:

$$\text{logit}(y)=\ln(\text{odds})=\ln\left(\frac{p}{1-p}\right)=a+\beta x \ldots\ldots (3)$$

Where *p* is the probability of interested outcome and *x* is the explanatory variable. The parameters of the logistic regression are α and β. This is the simple logistic model.

Taking the antilog of equation 1 on both sides, one can derive an equation for the prediction of the probability of the occurrence of interested outcome as

$$P = \frac{e^{\alpha+XB}}{1+e^{\alpha+XB}} = \frac{1}{1+e^{\alpha+XB}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

5.2 Odds ratio

It is a measure of effect size, describing the strength of association 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. An odds ratio is the ratio of two odds, such as the ratio of the odds for men and the odds for women. Here $p/(1-p)$ measures the probability that $y = 1$ relative to the probability that $y = 0$ and is called the **odds ratio** or **relative risk**. Odds ratios are the main effect size measure for logistic regression, reflecting in this case what difference gender makes as a predictor of some dependent variable. An odds ratio of 1.0 (which is 1:1 odds) indicates the variable has no effect, the furtherfrom1.0 in either direction, the greater the effect (Mbugua, M.D., 2014).

Extending the logic of the simple logistic regression to multiple predictors, one may construct a complex logistic regression as

$$\text{Logit}(y)=\ln\left(\frac{p}{1-p}\right)=\alpha+B_1X_i+B_2X_{2i}+\ldots+B_kX_{ki} \quad \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Therefore,

$$P=\frac{e^{a+B_1X_1+B_2X_2+\ldots+B_kX_k}}{1+e^{a+B_1X_1+B_2X_2+\ldots+B_kX_k}}=\frac{1}{1+e^{-(a+B_1X_1+B_2X_2+\ldots+B_kX_k)}} \quad \ldots\ldots\ldots (6)$$

Where p probability that a case is in a particular category, e base of natural logarithms (approx 2.72), α constant of the equation and, B coefficient of the predictor variables.

This ought to look somewhat similar to the log odds equation. The odds ratio for a particular predictor variable is defined as e^{β}, where ß is the logit coefficient estimate for the predictor and *e* is the natural log.

If ß is zero, the odds ratio will equal 1 (i.e., since any number to the 0 power is 1), which leaves the odds unchanged. If ß is positive, the odds ratio will be greater than 1, which means the odds are increased. If ß is negative, the coefficient will be less than 1, which means the odds are decreased (Ron Heck, 2012).

5.3 Estimating the Model Parameters

According to Mbugua, M.D.,(2014). In this section discussed how model parameters are estimated using the method of maximum likelihood and assessment of the fitted model using a Wald x^2 statistic and the likelihood ratio test. The likelihood for a given model is interpreted as the joint probability of the observed outcomes expressed as a function of the chosen regression model (Dietz et al., 2005). The model coefficients are unknown quantities and are estimated by maximizing their probabilities. It is useful when investigating the contribution of more than one predictor, or for predictors with multiple levels.

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

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

Substituting in using logistic regression model:

$$\ln L = \sum_i y_i \beta x_i - \sum_i \ln(1 + \exp(\beta x_i)) \quad (8)$$

5.4 Fitting the Logistic Regression Model

After estimating the regression coefficients, it is necessary to assess the appropriateness, adequacy, and usefulness of the model. First, the importance of each of the explanatory variables is assessed by carrying out likelihood ratio test or a

Wald x^2 statistic. The overall goodness of fit of the model is then tested.

5.4.1 The likelihood ratio test

The likelihood-ratio test used to assess overall model fit can also be used to assess the contribution of individual predictors to a given model.

The likelihood ratio test for a particular parameter compares the likelihood of obtaining the data when the parameter is zero ($L0$) with the likelihood ($L1$) of obtaining the data evaluated at the MLE of the parameter. The test statistic is calculated as follows:

G= -2 ln L0 = -2 (ln L0 -ln L1) L1 .... (9)

This statistics is compared with an x^2distribution with 1 degree of freedom.

To assess the contribution of individual predictors one can enter the predictors hierarchically, and then compare each new model with the previous model to determine the contribution of each predictor.

5.4.2 The Wald Statistic

Wald x^2statistic is used to test the significance of the individual regression coefficients in the model.

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

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

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 x^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.

5.5 Goodness of Fit of the Model

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.

5.5.1 Cox-SnellR^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^2statistics have been developed.

The *Cox-Snell* R^2 is a pseudo - R^2statistic, 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 (HosmerJr, Lemeshow, and Sturdivant, 2013). It is given by:

Cox-snell$R^2 = 1-\left[\frac{L(R)}{L(F)}\right]^{2/N}$ ……………………………………….(11)

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

5.5.2 NagelkerkeR^2

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

Nagelkerke $R^2 = \frac{1-\left[\frac{L(R)}{L(F)}\right]^{2/N}}{1-L(R)^{2/N}}$ …………………………………..(12)

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

5.5.3 Chi-Square Tests / Omnibus Tests

With logistic regression, instead of R^2 as the statistics for overall fit of the linear regression model, deviance between observed values from the expected values is used. In addition, Omnibus test as a general name refers to an overall or a global test; other names include F-test or Chi-squared test. In linear regression, residuals can be defined as $y_i - \hat{y}_i$, where y_i is the observed dependent variable for the i_{th} subject and $\hat{y}_i$ the corresponding prediction from the model. The same concept applies to logistic regression, where y_i is equal to either 1 or 0, and the corresponding prediction from the model is as

$$\hat{y}_i = \frac{\exp(\alpha+\beta_1 X_{i1}+\cdots\ldots+\beta_k X_{ik})}{1+\exp(\alpha+\beta_1 X_{i1}+\cdots\ldots+\beta_k X_{ik})} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(12)$$

Chi-square test can be based on the residuals, $y_i - \hat{y}_i$ (Peng & So, 2002). A standardized residual can be defined as rir2

$$r_i = \frac{y_i-\hat{y}_i}{\sqrt{\hat{y}_i(1-\hat{y}_i)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots(13)$$

Where the standard deviation of the residuals is $\hat{y}_i$ $(1\text{-}\hat{y}i)$, one can then form an x^2 statistic as

$$x^2 = \sum_{i=1}^{n} r_i^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(14)$$

H_0: the model not significant

H_A: the model is significant

This statistic follows an x^2 distribution with $n-(k+1)$ degrees of freedom so that p-values can be calculated. We reject the null hypothesis if the p-value < 0.05 that means the model is significant and represents the data well.

5.5.4 Hosmer-Lemeshow Test

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. The test statistic is calculated using the observed and expected counts for the categories as

$$H = \sum_{g=1}^{10} \frac{(Og - Eg)^2}{Eg} \quad \text{...} \quad (15)$$

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).

5.6 Enter

Enter a procedure for variable selection in which all variables in a block are entered in a single step.

6. RESULTS AND DISCUSSIONS

The logistic regression techniques have been applied to evaluate the social characteristics of the household's head as the determinants of household poverty in Sudan.

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 social variables

N	Explanatory Variable	Abbreviations	Definition	Characteristic
1	Can read and write with understanding	CRW	Refers to know the heads read and write with understanding	Yes = 0, No = 1
2	Level education of household head	LEHH	refers to the highest level of schooling that a person has reached	Category: 1-no qualification 2- primary 3- intermediate 4- secondary 5- university 6- post graduate 7-khalowa1*
3	Vocational training/ craft Of household head	VT	Training that emphasizes knowledge and skills needed for a specific trade, craft or job function.	Yes =0, No = 1
4	Marital status of household head	MS	The term describes whether the head of household is married or not.	Category: 1- never married 2- married 3- Widow 4- divorce
5	Disability type of household head	Dis	Suffering from any type of disability that prevents from doing usual work	Yes = 1, No = 0

1= (informal education to read holy Qur'an)*

Estimation binary logistic regression analysis for social variables

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

Variable	Pearson Chi-Square		Likelihood Ratio		df
	Value χ^2	P-value	Value χ^2	P-value	
CRW	71.320	.000	70.150	.000	2
HLS	201.891	.000	225.771	.000	7
VT	73.015	.000	72.540	.000	2
MS	36.685	.000	35.536	.000	4
Dis	4.442	.108	4.857	.088	2

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 are a very strong association between can read and write with understanding (CRW), the highest level of school ever completed (HLE), ever attended vocational training (VT) and marital status (MS) with poverty household status, except for the variable type of disability (Dis) was not significant at 5% level of significance, but it's significant when included all variables in the logistic regression. Also, can read and write with understanding not significant when analysis with the logistic regression.

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

							95% C.I. for EXP(B)	
	B	**S.E**	**Wald**	**df**	**Sig**	**Exp(B)**	**Lower**	**Upper**
Step1 CRW			4.197	2	.123			
CRW(1)	-1.521	.821	3.436	1	.064	.219	.044	1.091
CRW(2)	-1.608	.822	3.830	1	.050	.200	.040	1.002
HLS			125.05	7	.000			
No qu(1)	.018	.202	.008	1	.930	1.018	.686	1.511
Prim(2)	-.283	.205	1.903	1	.168	.753	.504	1.126
Inter(3)	-.274	.248	1.220	1	.269	.761	.468	1.236
Sco(4)	-.728	.211	11.852	1	.001	.483	.319	.731
Univ(5)	-1.800	.290	38.448	1	.000	.165	.094	.292
High(6)	-1.415	.639	4.913	1	.027	.243	.069	.849
Khal(7)	.241	.203	1.402	1	.236	1.272	.854	1.895
VT			4.034	2	.133			
VT(1)	-.169	.117	2.079	1	.149	.844	.671	1.063
VT(2)	-.276	.178	2.409	1	.121	.759	.536	1.075
MS			34.709	4	.000			
Never(1)	-.142	.810	.031	1	.861	.868	.178	4.241
Mar(2)	-1.195	.259	21.308	1	.000	.303	.182	.503
Wid(3)	-.770	.179	18.523	1	.000	.463	.326	.657
Div(4)	-.413	.205	4.060	1	.044	.662	.443	.989
Dis			9.259	2	.010			
Dis(1)	-20.69	40192.9	.000	1	1.00	.000	.000	.
Dis(2)	-.394	.129	9.259	1	.002	.674	.523	.869
Constant	1.950	.838	5.410	1	.020	7.029		

Source: Prepared by researcher from the Survey Data, 2015

Variable(s) entered on step 1: , CRW, HLS, VT, MS and Diss. a

According to the summary statistics from the table (3) are showing that the odds ratios of all variables are less than one that puts all variables in negative relation with the poverty status, except two variables (no qualification and khalowa).

The overall result of the table above shows that out of the fifth identified variables only two variables were not significant in explaining whether household's status is poor or not poor. The study shows that the variable can read and write with understanding does not a significantly affect poverty level probably (Wald=4.187, P-value=.123) and negative relation with the probability of being poor, because of lower education have more chances to be poor. The can read and write with understanding (illiteracy) increases the chance of being poor with 20.9% but is not significant influencing poverty, because more than 62% of the population of the school age knows how to read and write.

Generally, the results depict that there was a negative relationship between probability of being a poor and different level of education. It means that higher levels of education reduce the probability of being poor gradually. If we look at the no qualification, primary, intermediate and khalowa say are not statistically significant, All of them can be classified as weak qualification, because as the primary and intermediate stages have been integrated into one stage since 1990 and There has been a change in the educational ladder, so they have no any impact to poverty. But the secondary, university and higher education remain an important determinant of household welfare and say are statistically significant variable and are negative coefficients indicate that increased education has a significant impact in reducing the probability of being poor, implying that a higher level of education provides greater opportunities for a better job and, subsequently, a higher income. This implies that education is the important factors in reducing the impact of poverty at the household level. These findings confirmed the conclusions of other studies, such as Bigsten et al. (2003); Achia, (2010), Sarwar et al (2012) and Xhafaj, & Nurja, (2014). The results also indicate that the vocational training (Wald=4.034, P-value=.133) is not significant in explaining the probability of being a poor and negative relation with the probability of being a poor, this was due to lack of adequate training in Sudan.

Furthermore, the marital status is a statistically significant variable (Wald =34.709, P-value=.000) and negatively correlated to responsiveness. Moreover, the study shows that the married household heads have a higher chance of being poor as compared to household heads that are not married. More specifically, the results indicate that the married, widowed and divorced head of households, were significantly more likely to be poor than their never married. This may be as a result of having more dependants depending on the household head. We also found that the never married is a statistically insignificant, those who lived together, and enjoy lower welfare with our families and not married yet, for this they have not impact on poverty. Moreover, we found that the Suffering type of disability is significant (Wald = 9.259, P-value=.010) in explaining the probability of being a poor and negative relation with poverty.

Table (4) Omnibus Tests of Model Coefficients

	chi-square	df	sig
Step 1 Step	278.923	17	.000
Block	278.923	17	.000
Model	278.923	17	.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=278.923, df = 17, 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 of the model with selected covariates

Summary Statistic test	Value	Df	P-value
Hosmer - Lemeshow	4.436	7	.728
LR chi-square	84.70	17	.000
Log likelihood	-3692.917		
Cox and Snell R^2	.46		
Nagelkerke R^2	.64		
Pseudo R^2	.0113		

Source: Prepared by researcher from the Survey Data, 2015

According to the table (5), the variables are significant predictors of poverty ($p < \mathbf{0.05}$) the 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 $\boldsymbol{\chi^2}$of 4.436 and was not significant Suggesting that there was goodness of fit the model of the data. Thus we accept the null hypothesis that household characteristics and perceptions have an influence on poverty. The log likelihood yields a $\boldsymbol{\chi^2}$ of -3692.917 and was significant at ($p < \mathbf{0.05}$) which 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 0.46 and 0.64 which indicate that the model explains 46% to 64% of the variance in the outcome respectively. This good 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 0.114 which means that 11.3% 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	4015	45	98.9
	poor	1841	63	3.3
Overall Percentage				68.4

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

The regression classification table (6) revealed that binary logistic model managed to predict 68.4% of the responses correctly. Apart from percent correct predictions, the model Chi-Square statistic have been run to evaluate the performance of the model. Accordingly, the Chi-Square value was found to be 4.436 and the overall model was found significant at 0.05 levels. Moreover, the results indicate that, the model makes a correct prediction of the cases compared to 68.1% in the null model, a marked improvement. Of the 4060 households with non-poor, the model correctly identified 4015 of them as not likely to have one. Similarly, of 1904 that did have a poor, the model correctly identifies 63 as likely to have one. 3.3% is also known as the **sensitivity** of prediction. 98.9% 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= 1.950 - 1.521 CRW (1) -1.608 CRW (2) + .018 No qu - .283 Prim - .274 Intr - .728 Sco - 1.800 Univ -1.415 High + .241 Khl - .169 VT (1) - .276 VT (2) - .142 Nver - 1.195 Mar - .770 Wido - .413 Div - 20.69 Dis (1) - .394 Dis (2)……………………………………………….. (16)

The model indicates that out of the many variables identified as possible determinants of poverty status only seven were statistically significant.

They include; secondary (4), University (5), higher level (6), married (2), widowed (3), divorce (4) disability (2).

$$P = \frac{e^{a+B_1X_1+B_2X_2+\dots+B_kX_k}}{1+e^{a+B_1X_1+B_2X_2+\dots+B_kX_k}}, \quad 0 < p < 1 \quad \dots\dots\dots\dots\dots\dots \quad (17)$$

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 68.4% of the observations.

Furthermore, the highest level of school ever completed, marital status and Suffering from any type of disability are statistically significant and the sign on the parameter estimate support expectations, while the can read and write with understanding, ever attended vocational training are not significant.

7. CONCLUSIONS AND RECOMMENDATIONS

7.1 Conclusions

Therefore we can conclude from the results reported above that:

- The result reveals that highest levels of school, marital status, and type of disability significantly explains the poverty status of a household.
- No qualification, primary, intermediate and khalowa are not statistically significant, Read and write with understanding and vocational training had the insignificant impact on poverty.
- The highest level of school, marital status, type of disability, read and write and vocational training was positively associated with the probability of being poor.

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.
- ✓ This study recommends a careful review on the reforms to be taken in relation to education and poverty, suggests that expansion of education and vocational training programmed at the grass root level are amongst rural areas.
- ✓ The government should be focusing on improving the livelihood, education situation and health services.
- ✓ The results presented in this paper suggest that the NHBPS data can be used to determine the correlates of poverty with different variables.
- ✓ Additionally, similar studies can also be done using other poverty estimation techniques to be able to compare the results.

REFERENCES

[1] Achia, T.N.O., Wangombe, A. & Khadioli, N. 2010. A logistic regression model to identify key determinants of poverty using demographic and health survey data. European Journal of Social Sciences, 13(1), 38-45.

[2] Afzal, M., Malik, M.E., Begum, I., Sarwar, K. and Fatima, H., 2012. Relationship among education, poverty and economic growth in Pakistan: an econometric analysis. Journal of Elementary Education, 22(1), pp.23-45.

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

[4] Bigsten, A., Kebede, B., Shimeles, A. and Taddesse, M., 2003. Growth and poverty reduction in Ethiopia: Evidence from household panel surveys. World development, 31(1), pp.87-106.

[5] Castaneda, R.A., Doan, D., Newhouse, D.L., Nguyen, M.C., Uematsu, H. and Azevedo, J.P., 2016. Who are the Poor in the Developing World?.

[6] Central Bureau of Statistics, poverty in northern Sudan, Household Survey data, 1992, Khartoum.

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

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

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

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

[11] Lee, L.F., 1983. Generalized econometric models with selectivity. Econometrica: Journal of the Econometric Society, pp.507-512.

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

[13] Myftaraj, E., Zyka, E, and Bici, R., 2014. Identifying household-level determinants of poverty in Albania using logistic regression model. International Journal of Sustainable Development, 7(3), pp.35-42.

[14] Park, H., 2013. An introduction to logistic regression: from basic concepts to interpretation with particular attention to the nursing domain. Journal of Korean Academy of Nursing, 43(2), pp.154-164.

[15] Peng, C.Y.J., Lee, K.L. and Ingersoll, G.M., 2002. An introduction to logistic regression analysis and reporting. The journal of educational research, 96(1), pp.3-14

[16] Ron Heck, S; 2012. Seminars; Introducing Generalized Linear Models: Logistic Regression. Application Seminar.

[17] World Bank; 2000. World Development Report 2000/2001: Attacking Poverty, New York: Oxford University Press.

[18] Xhafaj, E. and Nurja, I., 2014. Determination of The Key Factors That Influence Poverty Through Econometric Models. European Scientific Journal, ESJ, 10(24).

[19] Yusyf, H. M., 2015. The determinants of Rural Poverty in Tanzania.

1- https://borgenproject.org/causes-of-poverty-in-sudan/ Wednesday.23.5.2018

2- borgenproject.org/why-is-Sudan-poor/ the Blog-Thursday – may-2018

3- https://borgenproject.org/poverty-sudan.Wednesday.23.5.2018

4- http://www.worldbank.org/en/country/sudan/overview Wednesday.23.5.2018

5- https://theodora.com/wfbcurrent/sudan/sudan_economy.htmll