# Determinants of Rural Households' Poverty with Selected Link Functions: The Case of Soro District, Hadiya Zone, SNNPR States, Ethiopia

**Denano Tesfaye[1]\* and Tekle Getachew[2]**

[1]Department of Business Administration, Dilla University, Ethiopia
[2]Department of Economics, Wachemo University, Ethiopia

## Abstract

Poverty is a general feature in Ethiopia causing many sufferings to the largest proportion of the population. This study assessed the determinants of rural household poverty in five selected kebeles in Soro district by using the generalized linear modeling approach. With the specific objectives of estimating rural households' poverty status, identifying appropriate link function, and identifying determinants of rural households' poverty. Primary data were obtained through structured questionnaire interview. A total of 184 selected sample households were identified by proportional allocation. Based on the primary data whereby set of consumption food energy-intake method the probability of being poor was assessed. By using national poverty line of 2200 kcal, out of total of 184 sample households surveyed 65.76% were found to be poor. Log-log link function is found to be more appropriate to fit the data. Model adequacy diagnostic tests of the cook's distance and GLM residuals shows that there were no outliers and influential values that had significant impact on the model results. Based on generalized linear model results, the major determinants of rural households' poverty were age of household heads, family size per adult equivalent, access to credit service, dependency ratio, TLU, access to health service, and number of oxen ownership. Hence, promoting equitable economic growth, family planning, increasing land productivity, increasing credit service, increasing health service, increasing TLU and promoting research extension farmer linkage are indispensable policy interventions to better reduce rural poverty.

**Keywords**: GLMs; Link functions; Rural poverty; Determinants; Soro district

## Introduction

Poverty in Ethiopia is a longstanding problem affecting a significant portion of its rural and urban population. Survey results of HICES indicated that the proportion of population below poverty line in Ethiopia stood at 30.4% in rural areas and 25.7% in urban areas in the 2010 fiscal year [1]. Recently, Ethiopia ranked 173rd out of 187 countries in its HDI value of 0.396 and the country's MPI value was 0.564 [2]. It is pervasive and widespread in Ethiopia [3]. The government statistics shows that 29.6 per cent of the total population of the country lives below the national poverty line. Moreover, poverty is more prevalent in rural (30.4 per cent) than urban areas 25.7 per cent [4]. Other studies also confirm that poverty disproportionately affects people in the rural areas [4,5].

The causes of rural poverty are many including wide fluctuations in agricultural production as a result of drought, ineffective and inefficient agricultural marketing system, under developed transport and communication networks, underdeveloped production technologies, limited access of rural households to support services, environmental degradation and lack of participation by rural poor people in decisions that affect their livelihoods. However, the persistent fluctuation in the amount and distribution of rainfall is considered as a major factor in rural poverty. Poor people in rural areas face an acute lack of basic social and economic infrastructure such as health and educational facilities, veterinary services and access to safe drinking water. Households headed by women are particularly vulnerable. Women are much less likely than men to receive an education or health benefits, or to have a voice in decisions affecting their lives. For them, poverty means high numbers of infant deaths, undernourished families, lack of education for children and other deprivations [6].

The important part in most poverty analysis is the identification of the poor from non-poor, which necessitates the poverty line to be determined given the appropriate measure of welfare. Poverty line is understood as a level of standard of living below which a household is considered as being in poverty. There are a number of methods to determine the poverty line. The most commonly applied methods are the direct calorie intake, the Food-Energy Intake and the Cost-of-Basic-Needs methods.

In this study, the Food-Energy Intake method is used based on three premises. First, since the prices of food items have shown a drastic rise in the past couple of years nationwide in general and in the study area in particular, the consumption expenditure of the households may not reflect their true consumption habit thereof. Consequently, consumption expenditure might inflate the result and hardly show the contemporary reality. Secondly, since in most of the cases, it is not uncommon to see households to understate their income and overstate their expenditure. This may also end up in distorted results which may make the analysis unreliable. Finally, since using the CBN method needs enumeration and quantification of the basics and non-basic needs in monetary terms, it may be difficult to the households enumerate and quantify correctly. In lieu of this, the study used the FEI method to delineate the poor from the non-poor.

Therefore, a basket of food items that yields a stipulated minimum

energy requirement of 2,200 kilo calories of energy per person per day as stipulated by the World Health Organization (WHO) was first estimated for households. This basket of goods was borrowed from earlier studies on urban poverty in Ethiopia. These studies established the food basket by first estimating the average quantities of the various food-items most frequently consumed by households in the lower 50 percent of per capita consumption expenditure and adjusting the calorie content of the different food items to yield the minimum stipulated calorie.

Generalized linear model attempts like other regression models, to fit a regression line though the data minimizing the sum of squares with the help of link functions [7]. Generalized linear model can handle many different distributions of a link function has to introduced. In most regular linear regression models the data is normally distributed to the mean of the distribution falls on the regression line. Generalized linear models include three components: 1) a random component, which is the response and an associated probability distribution; 2) a systematic component, which includes explanatory variables and relationships among them and 3) a link function, which specifies the relationship between the systematic component or linear predictor and the mean of the response. It is the link function that allows generalization of the linear models for count, binomial and percent data thus ensuring linearity and constraining the predictions to be within a range of possible values. The link function is a way to link together the mean of different distributions within the generalized linear model to minimize the sum of squares of the data.

The fact that most studies on poverty determinants in Ethiopia have concentrated on logistic regression methods makes this study to be very essential as it allows the application of generalized linear models, based on link functions as well as an assessment of the performance of link functions.

## Methodology

### Generalized models

Generalized linear models (GLMs) are a large class of statistical models for relating responses to linear combinations of predictor variables, including many commonly encountered types of dependent variables and error structures as special cases. In addition to regression models for continuous dependent variables, models for rates and proportions, binary, ordinal and multinomial variables and counts can be handled as GLMs. The GLM approach is attractive because it;

1. Provides a general theoretical framework for many commonly encountered statistical models;

2. Simplifies the implementation of these different models in statistical software, since essentially the same algorithm can be used for estimation, inference and assessing model adequacy for all GLMs.

In this study, we consider generalized linear modeling approach in which the outcome variables are measured on a binary scale. For example, the response is poor or non-poor.

First, we define the binary random variable:

$$y_i = \begin{cases} 1, \text{if the household is poor} \\ 0, \text{if the household is non-poor} \end{cases} \quad (1)$$

With probabilities $pr(y_i=1/x)=$ and $pr(y_i=0/x)=1-\pi$ which is the Binomial distribution. Binomial distribution is a part of generalized linear models. The generalized linear models were formulated [6] and

discuss estimation of the parameters. Let $y_1,...,y_n$ denote k independent observations on a response. We make $y_i$ as recognition of a random variable $y_i$. In the general linear model (GLM) we assume that $y_i$ has a normal distribution with mean $\mu_i$ and variance $\delta^2$ i.e $y_i \sim N(\mu, \delta^2)$, and we further assume that the expected value $\mu_i$ is a linear function of k predictors the take value $x_{i1}=x_{i1},...., x_{ik}$ for the $i^{th}$ case, so that $\mu_i=x_i\beta$, where $\beta$ is a vector of unknown parameters.

Generally, three components of generalized linear models:

1. The random component: the component of Y have independent normal distributions with E(Y)=$\mu$ and constant variance $\delta^2$;

The systematic component: covariates $x_1,....x_p$ produce a linear predictor $\eta$ given by $\eta = \sum_{i=1}^{p} x_i \beta_i$ , and

The link between the random and systematic components: $\mu = \eta$.

This generalization introduces a new symbol $\eta$ for the linear predictor and the third components are specifies that $\mu$ and $\eta$ are in fact identical. If we write, $\eta_i = g(\mu_i)$: then, g(.) will be called the link function. A generalized linear model is defined by specifying two components. The response should be a member of the exponential family distribution and the link function describes how the mean of the response and a linear combination of the predictors are related.

### Exponential family

Recall the representation [8] the exponential family just defined includes as special cases the normal, binomial, Poisson, gamma, beta, geometric, negative binomial and inverse Gaussian distributions. In a generalized linear model, the distribution of response variable is from the exponential family of distributions which take the general form:

$$f(y/\theta,\varphi) = \exp\left[\frac{y\theta - b(\theta)}{a(\varphi)}\right] + c(y,\varphi) \quad (2)$$

The $\theta$ is called the *canonical* parameter and represents the location while $\phi$ is called the dispersion parameter and represents the scale. We may define various members of the family by specifying the functions *a, b,* and *c*.

For this study, we use binomial function:

$$f(y/\theta,\varphi) = \binom{n}{y}\mu^y(1-\mu)^{n-y} \quad (3)$$

$$= \exp\left(y\log\mu + (n-y)\log(1-\mu) + \log\binom{n}{y}\right)$$

$$= \exp\left(y\log\left(\frac{\mu}{1-\mu}\right) + n\log(1-\mu) + \log\binom{n}{y}\right)$$

Where, $\theta = \log\left(\frac{\mu}{1-\mu}\right)$, $b(\theta) = n\log(1+\exp\theta)$ and $c(y,\varphi) = \log\binom{n}{y}$

Therefore, the binomial distribution is an exponential family distribution.

Why because, generalized linear model is a distribution of response variable from the exponential family of distributions.

The likelihood function for estimation of exponential distribution is

$$L(\theta,\varphi;y) = \prod_{i=1}^{n} f y(y_i;\theta_i,\varphi) = \prod_{i=1}^{n} \exp\left[\frac{y\theta - b(\theta)}{a(\varphi)}\right] + c(y,\varphi) \quad (4)$$

It then, follows that the log-likelihood function is

$$l(\theta,\varphi;y) = \ln L(\theta,\varphi;y) = \sum_{i=1}^{n} fy(y_i,\theta_i,\varphi)$$

$$l(\theta,\varphi;y) = \ln(\sum_{i=1}^{n} \frac{y_i\theta_i - b(\theta_i)}{a(\varphi)} + c(y_i,\varphi))$$

$$l(\theta,\varphi;y) = \ln(\sum_{i=1}^{n} \left\{ \frac{y_i\theta_i - b(\theta_i)}{a(\varphi)} \right\} + \sum_{i=1}^{n} c(y_i,\varphi))$$

Therefore, taking the derivation of $l(\theta,\varphi;y)$ with the respect to the $\theta_i's$, the dispersion term $\varphi$ factors out for some $a(\varphi)$ functions so that the estimation of the canonical parameters $\theta_i$ can be carried out separately from that for the parameter $\varphi$.

## Link functions

The link function is a way to link together the mean of the distributions within the GLMs to minimize the sum of squares of the data. Link function is therefore a way together a dependent variable that is non-linear to the other variables being evaluated. The reason for this is based on the fact that the response variable is dichotomous (0=non-poor and 1=poor). There are four link functions considered were Logit, Log-log, Probit and Complementary log-log link functions respectively. The performance of each selected link function was assessed based on the value of its deviance statistic and Akaike information criteria. The generalized linear models yielding the minimum deviance statistic and Akaike information criteria were adjudged best in all cases.

## Complementary log-log link function

Under the assumption of binary response, there are three alternatives to logit model: probit model, log-log model and complementary log-log model.

The form of Complementary log-log model is:

$$\pi_i = 1 - \exp(-\exp(x^T\beta)) \tag{5}$$

$$\eta = \log(-\log(1-\pi_i)) = x^T\beta$$

Where, $\eta$ is the complementary log-log link function that associate the outcome variable with the predictor variables a linear relation, x is nxp matrices of predictors, $\beta$ is px1 vectors of coefficients for the predictors and $\pi_i$ is defined as the success the probability corresponding to the $i^{th}$ observations.

## Log-log link function

When we use the probability of success in place of the probability of failure in the complementary log-log model the model become log-log model.

The model becomes:

$$\prod_i = -\exp(-\exp(-X^T\beta)) \tag{6}$$

$$\eta = \log(-\log(\prod_i)) = \mathbf{X}^T\beta$$

Where, $\eta$ is the log-log link function that associate the outcome variable with the predictor variables a linear relation; x is nxp matrix of predictors; $\beta$ is px1 vectors of coefficients for the predictors and $\pi_i$ is defined as the success the probability corresponding to the $i^{th}$ observations.

## Logit link function

Logistic regression allows the prediction of a discrete outcome such as group membership from a set of explanatory variables that may be continuous, discrete, and dummy or a mixture of these. It is one in which the dependent variable is binary and assumes only two values like absence or presence and poor or non-poor. In this study, the presence of the outcome of interest, say being poor is assigned the value of 1, while the absence of the outcome of interest, say not being poor is assigned the value of 0.

Logistic link function written as,

$$\prod_i = \left( \frac{\exp(x^T\beta)}{1+\exp(x^T\beta)} \right) \tag{7}$$

$$\eta = \log\left( \frac{\prod_i}{1-\prod_i} \right) = \mathbf{X}^T\beta$$

## The probit link function

The idea of probit analysis was originally published in *Science* by Chester Ittner Bliss. He worked as an entomologist for the Connecticut agricultural experiment station and was primarily concerned with finding an effective pesticide to control insects that fed on grape leaves [9]. By plotting the response of the insects to various concentrations of pesticides, he could visually see that each pesticide affected the insects at different concentrations, i.e. one was more effective than the other. However, he didn't have a statistically sound method to compare this difference. In 1952, a professor of statistics at the University of Edinburgh by the name of David Finney took Bliss' idea and wrote a book called Probit Analysis [10].

Household consumption is below the estimated poverty line, the household is considered as poor, otherwise zero. Poverty line is established based on the estimated amount of monetary value that is required to meet the basic needs of the household for a month. If the household is poor it takes the value 1 otherwise zero. Then, the predicted values of the dependent variable lie on zero and one. Hence, the predicted values are interpreted as probabilities.

The model is written as:

$$\prod_i = \int_{-\infty}^{\mathbf{X}^T\beta} \varphi(t)dt = \int_{-\infty}^{\mathbf{X}^T\beta} \frac{1}{\sqrt{2\prod\delta^2}} \exp\left(-\frac{1}{2}t^2\right)dt \tag{8}$$

$$\eta = \varphi^{-1}(\prod_i) = \mathbf{X}^T\beta$$

Where, $\eta$ is the probit link function that associate the outcome variable with the predictor variables a linear relation, x is nxp matrix of predictors, $\beta$ is px1 vectors of coefficients for the predictors, $\varphi$ is cumulative function and $\pi_i$ is defined as the success the probability corresponding to the $i^{th}$ observations.

## Maximum likelihood estimation for probit link functions

$$f(y/x,\beta) = \prod_{i=1}^{n} F(x'_i\beta)^{y_i} \left[1 - F(x'_i\beta)\right]^{1-y_i} \tag{9}$$

Let $F_i = x'_i\beta$

$$L = \prod_{i=1}^{n} F_i^{y_i} (1-F_i)^{1-y_i}$$

Log likelihood function

$$\ln L = \sum_i y_i \ln F_i + \sum_{i=1} (1-y_i)\ln(1-F_i)$$

$$\frac{\partial \ln L}{\partial \beta} = 0$$

## Model adequacy checking

In general linear models, the fit of the model to data can be explored by using residual plots and other diagnostic tools. The purpose of model diagnostics is to examine whether the model provides a reasonable approximation to the data. If there are indications of systematic deviations between data and model, the model should be modified.

**Hat matrix:** In general linear models, a residual is the difference between the observed value of y and the fitted value $\hat{y}$ by that would be obtained if the model were perfectly true $e = y - \hat{y}$. The concept of a "residuals" is not quite as clear-cut in generalized linear models. The estimated expected value ("fitted value") of the response in a general linear model is $E(Y_i) = \hat{\mu}_i = \hat{y}_i$. The fitted values are linear functions of the observed values. For linear predictors, it holds that:

$$\hat{y} = \mathbf{H}y \qquad (10)$$

Where, H is known as the "hat matrix", H is idempotent (i.e HH=H) and symmetric.

**Residuals in generalized linear models**: Residuals provide another way to assess the adequacy of a model. In general linear models, the observed residuals are simply the difference between the observed values of y and the values $\hat{y}$ that are predicted from the model: $\hat{e} = y - \hat{y}$. In generalized linear models, the variance of the residuals is often related to the size of $\hat{y}$. Therefore, some kind of scaling mechanism is needed if we want to use the residuals for plots or other model diagnostics. For binomial distributions, diagnostic plots of residuals do not seem useful; the two line residual plot is difficult to interpret. Instead [6] concentrate on diagnostics that identify where the model fits poorly. Several suggestions have been made on how to achieve in this way.

**Pearson residuals:** The raw residual for an observation $y_i$ can be defined as $\hat{e} = y_i - \hat{y}_i$. The Pearson residual is the raw residual standardized with the standard deviation of the fitted value:

$$e_i = \frac{y_i - \hat{y}_i}{\sqrt{\hat{\mathrm{var}}(\hat{y}_i)}} \qquad (11)$$

For binomial model the Pearson residuals become $e_i = \frac{y_i - n_i \hat{p}_i}{\sqrt{n_i \hat{p}_i (1 - \hat{p}_i)}}$:

If the model holds, Pearson residuals can often be considered to be approximately normally distributed with a constant variance, in large samples. However, even when they are standardized with the standard error of $\hat{y}$. This is since we have standardized the residuals using estimated standard errors. Still, the standard errors of Pearson residuals can be estimated. It can be shown that adjusted Pearson residuals can be obtained as.

$$e_i = \frac{e_i, pearson}{\sqrt{1 - h_{ii}}} \qquad (12)$$

Where, $h_{ii}$ are diagonal elements from the hat matrix. The adjusted Pearson residuals can often be considered to be standard Normal, which means that e.g. residuals outside ±3 will occur in about 5% of the cases. This can be used to detect possible outliers in the data.

**Deviance residuals:** Observation number i contributes an amount $d_i$ to the deviance, as a measure of fit of the model: $D = \sum_i d_i$. We define the deviance residuals as:

$$e_i = sign(y_i - \hat{y}_i)\sqrt{d_i} \qquad (13)$$

The deviance residuals can also be written in standardized form.

This is obtained as:

$$e_i, adj, D = \frac{e_i, Deviance}{\sqrt{1 - h_{ii}}}$$

Where, $h_{ii}$ are diagonal elements from the hat matrix.

**Likelihood residuals:** Theoretically it would be possible compare the deviance of a model that comprises all the data with the deviance of a model with observation i excluded. An approximation to the residuals that would be obtained using this procedure is

$$e_i, Likelihood = sign(y_i - \hat{y}_i)\sqrt{h_{ii}(e_i, score)^2 + (1 - h_{ii})(e_i, deviance)^2} \qquad (14)$$

Where, $h_{ii}$ are diagonal elements from the hat matrix.

## Model selection criteria

The Bayesian information criteria were introduced by Schwarz as a competitor to the Akaike information criterion. Schwarz derived BIC to serve as an asymptotic approximation to a transformation of the Bayesian posterior probability of a candidate model. The computation of BIC is based on the emperical log-likelihood. Akaike information criteria are also another method of model selection criteria. AIC is very important to identify an appropriate model for determinants of household poverty from binary choice models.

Akaike Information Criteria is given by:

AIC=-2 log (likelihood)+2 p  (15)

Bayesian Information Criteria is given by:

BIC=-2 log (likelihood)+p log n  (16)

Where, log (likelihood) is the log-likelihood function which measures the goodness of the fitted model. P is the number of parameter in the model which measures how complexity of fitted model; n is number of selected sample size in the model for the study. The minimum value of AIC or BIC was considered as an appropriate model for the study (Tables 1 and 2).

**Deviance:** If the response variables $y_{1,\,...,}\,y_n$ are independent and $y_1 \sim Bin(n_i, \pi_i)$, then the log-likelihood function is;

$$L(\beta, y) = \sum_{i=1}^{N}\left[ y_i \log \Pi_i - y_i \log(1 - \Pi_i) + n_i \log(1 - \Pi_i) + \log\binom{n_i}{y_i} \right] \qquad (17)$$

Where, $\beta = \left[ \Pi_1, ..., \Pi_N \right]^T$

The maximum likelihood estimates are $\Pi_i = \frac{y_i}{n_i}$, so the maximum value of the log-likelihood function is

$$L(\beta_{\max}, y) = \sum\left[ y_i \log\left(\frac{y_i}{n_i}\right) - y_i \log\left(\frac{n_i - y_i}{n}\right) + n_i \log\left(\frac{n_i - y_i}{n}\right) + \log\binom{n_i}{y_i} \right]$$

For any other model with P<N parameters, let $_i$ denote the maximum likelihood estimates for the probabilities and let $y_i = n_i \pi_i$ denote the fitted values. Then the log-likelihood function evaluated at these values is

$$L(\beta_j, y) = \sum\left[ y_i \log\left(\frac{\hat{y}_i}{n_i}\right) - y_i \log\left(\frac{n_i - \hat{y}_i}{n_i}\right) + n_i \log\left(\frac{n_i - \hat{y}_i}{n_i}\right) + \log\binom{n_i}{\hat{y}_i} \right]$$

Therefore, the deviance is

D=2[l($\beta_{max}$,y)-l($\beta_j$,y)]

$$D = 2\sum_{i=1}^{N}\left[ y_i \log\left(\frac{y_i}{\hat{y}}\right) + (n_i - y_i)\log\left(\frac{n_i - y_i}{n_i - \hat{y}_i}\right) \right]$$

| Types variable | Codes | Explanation of variables | Hypothesized effect on poverty |
|---|---|---|---|
| HHage | Continuous | Age of the household head in years | + |
| AE | Continuous | Family size in adult equivalent | + |
| Depratio | Continuous | Consumer-worker ratio of the household in percent | + |
| Farm-size | Continuous | Farm size in hectare of the household | - |
| Incomoff/non | Continuous | Household off-farm/non-farm income in Birr | - |
| TLU | Continuous | Household livestock number in Tropical Livestock Unit | - |
| Hhedu | Dummy | 1 for can read and write 0 otherwise | - |
| Hhsex | Dummy | Takes value of 1 if male and 0 otherwise | - |
| Dstmrkt | Continuous | Distance nearest to market in Kilometer | + |
| Accirrgtn | Dummy | Take value of 1 if Access to irrigation 0 otherwise | - |
| Oxen | Dummy | Take value 1 if access to Oxen of the household head | - |
| Acchtsv | Dummy | Take value 1 if access to health service, 0 otherwise | - |
| Credit | Dummy | Take value 1 if HHS received credit otherwise 0 | - |

**Table 1:** Types, codes and explanation of variables in the model.

| HH poverty status | Estimated coefficient | Std. error | Z-Statistics | Prob. |
|---|---|---|---|---|
| Hhsex | -0.2287 | 0.4963 | -0.46 | 0.645 |
| Age | 0.1061 | 0.0292 | 3.63 | 0.000*** |
| Hhedu | -0.2792 | 0.4578 | -0.61 | 0.542 |
| AE/family size | 0.6999 | 0.1918 | 3.65 | 0.000*** |
| Farm-size | -0.0504 | 0.299 | 0.17 | 0.866 |
| Dstmrkt | 0.0244 | 0.0193 | 1.26 | 0.206 |
| Accirrgn | -0.6443 | 0.6686 | -0.96 | 0.335 |
| Oxen | -1.2449 | 0.606 | -2.05 | 0.040** |
| Acchtsv | -1.1487 | 0.5858 | -1.96 | 0.050* |
| Depratio | 5.6383 | 1.617 | 3.49 | 0.000*** |
| TLU | -0.1468 | 0.0718 | -2.04 | 0.041** |
| Credit | -1.541 | 0.4697 | -3.28 | 0.001*** |
| Incomoff/non | -0.07968 | 0.4758 | -0.17 | 0.867 |
| Constant | -7.6895 | 3.0799 | -2.5 | 0.013 |

***, ** and * statistically significant at 1%, 5% and 10% respectively.

**Table 2:** Results for generalized linear modeling, family=Bernoulli.

## Results and Discussion

The descriptive statistics showed the existence of a significant mean difference in family size per adult equivalent, tropical unit livestock (TLU) and dependency ratio at less than one percent probability level between poor and non-poor households. The t-test for family size per adult equivalent, TLU and dependency ratio showed a mean difference between this entire variable at less than 1% probability level.

The sample households were classified into poor and non-poor groups on households' poverty status. As a result 2200 kcal per adult equivalent (AE) per day was employed as a cut-off point between poor and non-poor households'. Accordingly, the result of the study showed that about 65.74 percent of sample households were found to be poor.

In this study, the log-log link function performed best as it gave the minimum value of the deviance statistic and Akaike information criteria. The logit, probit and complementary log-log came second, third and fourth respectively. Therefore, the study indicated that log-log link function was more appropriate to identify the determinants of poverty in the households' level.

Model adequacy diagnostic tests of the cook's distance, Pearson residuals, deviance residuals and likelihood residuals shows that there were no outliers and influential values that had significant impact on the model results. Besides, the Collinearity diagnostic tests show that Multicollinearity was not a great threat to the reliability of model coefficients.

In the results of generalized linear model (GLM), each of thirteen predictors (sex, age, family size per adult equivalent, dependency ratio, TLU, oxen, access of credit, access of health service, access of irrigation, income from off-farm/non-farm, cultivated land size in hectare and number of oxen ownership) had the expected sign and were significant to be considered as candidates of the model. Seven out of thirteen predictor variables were selected as determinants of rural households' poverty. Based on generalized linear model results household head age, family size in adult equivalent, dependency ratio, access to credit service, TLU, access to health service and oxen ownership were the major determinants of rural households' poverty.

## Conclusion and Recommendations

Several study have shown the evidence that poverty is lack of capability, taking capability to mean to be able to live longer, to be well nourished, to be health, and to be literate, and the value of living standard lies in the living not in possession of commodities. Accordingly, the task of poverty analysis is to determine what those capabilities are in specific society and who fail to reach those deprivations.

Identifying and prioritizing the correlates of poverty is of principal importance in the endeavor to make supreme decisions in eradicating poverty and achieve the well-being of citizens. In an attempt to identify the poverty status on selected *kebeles* of *Soro district*, around 65% of survey households were found to be poor. These obviously entail the need for a clear picture of the variables determining poverty. Due to the multifaceted nature of poverty, it is absolutely absurd to say all the correlates that interplay is the area can be identified in simple terms. However, the one assumed to be most critical should be identified with their clear indication and brought into the attention of the concerned parties.

By looking critically for identifying the determinants of poverty status of rural households' in *Soro* district, the following are recommendations based on research findings.

1. The study has assessed only five *kebeles* of the district at household level. However, the researcher believes that more extensive surveys in other kebeles can generate additional results because of diverse demographic and socio-economic setups of the different community.

2. The livestock rearing are dual in purpose, for income and food sources. These alternative potentials of the husbandry have found great importance for family from getting poor. Projects like dairy cow credits, sheep and goat credits, and

fattening need to be supported with the husbandry skill and knowledge training for expansion on depth concepts livestock managements to increase family income. Managements of herds or stocking and restocking and utilization of improved feed and fodders need to be given due attention. Hence, the output of the livestock sector should be strengthened through the provision or supply of better veterinary services.

3. Access of credit services are the main sources of financial capital sustain rural livelihood. Thus, enhancing and expanding rural credit services to substance farmers in the district should be one of the primary areas of intervention and policy options. Rural credit service can help farm households in solving credit constraints faced to buy farm oxen, modern farm inputs, and enhance of technology adoption. Therefore, access of credit must be improved for poor rural households.

4. Household with large family size found difficult to secure basic necessities this because the family could have high dependency ratio. Thus the possibility of a family income to increase as additional family member added is very low or none. Projects/programs that work on family planning need to be encouraged maintaining and minimizing household size to the level of household income capability. With these scenario, having more household size aggravate the problem of meeting food leave alone education, health and other non-food demands of household that will bring future return. So, action based awareness creation on the impacts of population growth at the family, community and national level should be strongly advocated that lead to reduction in fertility and lengthen birth spacing resulted in smaller household size. Moreover, development actors involved on population issue should encourage households having acceptable number of children through provision of especial offer such as covering schooling cost, giving training and other related incentives.

5. Expansion of health institutions to more remote rural areas and building the capacity of the existing access of health services should be encouraged to improve the rural households' health status.

6. The log-log link function was more appropriate in fitting the determinants of rural household poverty in comparison generalized linear modeling approach with probit, complementary log-log and logit link functions.

7. All in all, the persistent problem of poverty in the study area can be controlled to a meaningful level if there is commitment on the side of different parties in identifying as well as prioritizing of the elements responsible for the incident and putting forward sound policies and actions in controlling them.

## References

1. MoFED (2012) Ethiopia's Progress towards Eradicating Poverty: An Interim Report on Poverty Analysis Study (2010/11): Ministry of Finance and Economic Development, Addis Ababa, Ethiopia.

2. Human Development Report (2013) The Rise of the South human progress: in a diverse world: Explanatory note on the 2013 HDR composite indices.

3. Christiaensen L, Demery L, Kuhl J (2010) The (Evolving) role of agriculture in poverty reduction: An empirical perspective. UNU-WIDER Working Paper: No. 2010/36.

4. Michael HK, Christopher J, Nachtsheim J (2005) Applied linear statistical models, (5thedn), Mc Graw Hill publications, USA.

5. MoFED (2012) Ethiopia's Progress towards Eradicating Poverty: An Interim Report on Poverty Analysis Study (2010/11): Ministry of Finance and Economic Development, Addis Ababa, Ethiopia.

6. Finney DJ (1952) Probit Analysis: Cambridge, England, Cambridge University Press, UK.

7. Greenberg BG (1980) Obituary Notices Chester I. Bliss, 1899-1979. International Statistical Review 8: 135-136.

8. IFAD (2010) Rural poverty report: New Realities, New Challenges: New Opportunities for Tomorrow's Generation International Fund for Agricultural Development (IFAD), Italy.

9. McCullagh P, Nelder JA (1989) Generalized Linear Models, (2ndedn), Chapman and Hall, UK.

10. World Bank (2008) World Development Report: Agriculture for Development, World Bank Institute.