On the Level of Performance of Selected Link Functions in the Identification of Poverty Correlates in Nigeria

Osowole, O. I.* Nwoke, Francisca, Uba, Ezenwanyi*****

*Department of Statistics, University of Ibadan, Nigeria
**National Bureau of Statistics, Ibadan, Oyo State, Nigeria
*** National Root Crops Research Institute, Umahia, Abia State, Nigeria
*Email of the Corresponding author: dosowole@yahoo.com

Abstract

Effective poverty reduction requires proper identification of correlates of poverty. This study therefore used the generalized linear modeling approach to identify poverty correlates in Nigeria based on the 2003/2004 National Living Standard Survey (NLSS) data. This approach is adequate because it respects the underlying distribution of poverty indicators as the response variable in models to facilitate inference on poverty. The results of the study show that sex of the household head, age in years of the household head, father's education level, father's work, mother's work, household size, occupation group of the household head, and educational group for highest level attained by the household head are strong correlates of poverty while mother's education level and literacy of the household head are weak correlates of poverty. Having fewer numbers of children and ensuring education of children are recommended as measures to reduce the burden of poverty.

Keywords: Link functions, Poverty correlates, Human dignity, Lack of understanding, Regression

1.0 Introduction

There is no general consensus on the definition of poverty. This is not unconnected to its multidimensional nature, which affects many aspects of human conditions, including physical, moral, social, and psychological aspects. Hence, many criteria have been used to define poverty. While an economist would approach the subject from the view point of wants, needs and effective demand, the psychologist may look at it from the standpoint of deprivation, esteem and ego. But whatever perspective it is viewed, it is obvious that it is a condition of life that is so degrading as to insult human dignity [1]. The poverty situation in Nigeria presents a paradox, because despite the human and material endowments of Nigeria, a large proportion of her Population is still poor. [2]

Using the most recent indicators of poverty such as illiteracy and access to safe water, Nigeria ranks below Kenya, Ghana and Zambia. The Nigeria's Gross National Product per capita is also lower while the purchasing power of Nigerians continues to decline with high inflation and increasing income inequality [3]. Using the 1980, 1985, 1992 and 1996 National Consumer Surveys' data revealed that poverty has been on the increase. With an approximate 17.7 million poor people in 1980, the number of poor people rose to 34.7 million in 1985, 39.2 million in 1992 and 67.1 million poor people in 1996.

With declining economic conditions within the economy, the population in poverty in 2001 would have increased far above 70 million people.

Apart from the very low growth in the economy and bad governance of the past, that have been responsible for the increasing poverty level of Nigerians, the lack of understanding of the causes of poverty empirically have made all efforts so far at poverty reduction to be unfelt. [4-5] have revealed that if effective policies to reduce poverty are to be formulated or proper poverty reduction programmes are to be established, more knowledge and understanding of the specific determinants of poverty are required. This is done by constructing the poverty profile in the form of a regression of the individual's poverty measure against a variety of household characteristics. In Nigeria, a lot of studies have been done on poverty [6-14].

A common method of regression according to [15] is to state that y_i/z_i (or its log) is a function of a vector of observed household characteristics x_i, namely $y_i/z_i = \beta x_i + e_i$ where β is a vector of parameters, e_i is the error term, y_i is the per capita expenditure, z_i is the poverty line and y_i/z_i is the welfare ratio. One then defines the binary variable $h_i = 1$ if $y_i/z_i < 1$and $h_i = 0$ otherwise. This method then pretends not to observe the y_i's, acting as if only h_i and the vector of characteristics x_i is observed. Hence, mostly logistic regression model is normally used because it has a specification which is designed to handle the specific requirements of a binary dependent variable model. The fact that most studies on poverty determinants in Nigeria have concentrated on classical and 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 selected in this study. The link functions considered are logit, log-log, complementary log-log and probit link functions. It is noteworthy that this study is perhaps the first in which different link functions are investigated simultaneously.

2,0 Data and Methods

The National Living Standard Survey (NLSS) 2004, institutionalized by the National Bureau of Statistics will be used in this study. This is because it provides a major survey framework for regular production, management and tracking of poverty programmes and policies. The survey was designed to give estimates at National, Zonal and State levels. The first stage was a cluster of housing units called Enumeration Area (EA), while the second stage was the housing units. One hundred and twenty EAs were selected and sensitized in each state, while sixty were selected in the Federal Capital Territory. Ten EAs with five housing units were studied per month. Thus a total of fifty housing units were canvassed per month in each state and twenty-five in Abuja. The National Bureau of Statistics (Nigeria) (NBS) field staff resident in the enumeration areas were responsible for data collection for the survey. They interviewed the households using a diary of daily consumption and expenditure to support the interviews. The NLSS data collected eventually were of good quality due to the effective supervision and quality control measures put in place by the National Bureau of Statistics [3]. For instance, a supervisor was attached to each team to observe interviews and confirm the pre-selected households. He was to verify and edit completed questionnaires. The state officers and zonal controllers conducted regular monitoring visits to the EAs. Headquarters monitoring groups also visited states on quarterly basis, for on-the-spot assessment of the quality of work. An independent firm was engaged to monitor the fieldwork in the states from the commencement to the end of the survey. Data were collected on the following key elements: demographic characteristics, educational skill and training, employment and time use, housing and housing conditions, social capital, agriculture, income, consumption expenditure and non-farm enterprise. Some of the variables captured in the survey included sector of the country, sex of the household head, age in years of the household head, marital status of the household head, religion of the household head, father's educational level, father's work, mother's educational level, mother's work, household size, expenditure of own produce, household expenditure on food, occupation group the household head belongs, educational group for highest level attained by the household, literacy of the household head and educational age grouping.

2.0.1 The Generalized Linear Model

We shall describe the generalized linear model as formulated by [16], and discuss estimation of the parameters. Let $y_1, \ldots, y_n$ denote n independent observations on a response. We treat y_i as a realization of a random variable Y_i. In the general linear model we assume that Y_i has a normal distribution with mean μ_i and variance σ^2, i.e., $Yi \sim N(\mu_i, \sigma^2)$, and we further assume that the expected value μ_i is a linear function of p predictors that take values $x_i = (x_{i1}, \ldots, x_{ip})$ for the i^{-th} case, so that $\mu_i = x_i\boldsymbol{\beta}$, where $\boldsymbol{\beta}$ is a vector of unknown parameters. We will generalize this in two steps, dealing with the stochastic and systematic components of the model.

2.0.2 The Exponential Family

We will assume that the observations come from a distribution in the exponential family with probability density function

$$f(y_i)=\exp\left(\frac{y_i\theta_i-b(\theta_i)}{a_i(\varphi)}+c(y_i,\varphi)\right) \quad (1)$$

Here θ_i and φ are parameters and $a_i(\varphi)$, $b(\theta_i)$ and $c(y_i, \varphi)$ are known functions. Generally, the function $a_i(\varphi)$ has the form $a_i(\varphi)=\frac{\varphi}{p_i}$ where p_i is a known prior weight, usually 1. The parameters θ_i and φ are essentially location and scale parameters. It can be shown that if Y_i has a distribution in the exponential family then it has mean and variance

$$E(Y_i)=\mu_i=b'(\theta_i) \quad (2)$$

$$Var(Y_i)=\sigma_i^2=b''(\theta_i)a_i(\varphi) \quad (3)$$

where $b'(\theta_i)$ and $b''(\theta_i)$ are the first and second derivatives of $b(\theta_i)$. When $a_i(\varphi) = \frac{\varphi}{p_i}$ the variance has the simpler form

$$Var(Y_i)=\sigma_i^2=\frac{\varphi b''(\theta_i)}{p_i} \quad (4)$$

The exponential family just defined includes as special cases the normal, binomial, Poisson, exponential, gamma, beta, geometric, negative binomial and inverse Gaussian distributions. For example the normal distribution has density:

$$f(y_i)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{(y_i-\mu_i)^2}{2\sigma^2}\} \quad (5)$$

Expanding the square in the exponent we get $(y_i-\mu_i)^2=y_i^2+\mu_i^2-2y_i\mu_i$ so the coefficient of y_i is $\frac{\mu_i}{\sigma^2}$. This result identifies θ_i as μ_i and φ as σ^2, with $a_i(\varphi)=\varphi$. Now write:

$$\left.\begin{array}{l} f(y_i)=\exp\{[\frac{y_i\mu_i-\frac{1}{2}\mu_i^2}{\sigma^2}-\frac{y_i^2}{2\sigma^2}-\frac{1}{2}\log(2\pi\sigma^2)]\}\text{where} \\ \theta_i=\mu_i, y_i\theta_i=y_i\mu_i, b(\theta_i)=b(\mu_i)=\frac{1}{2}\mu_i^2, b'(\mu_i)=\mu_i, b''(\mu_i)a_i(\varphi)=b''(\mu_i)\sigma^2=\sigma^2 \end{array}\right\} \quad (6)$$

This shows that $b(\theta_i)=\frac{1}{2}\theta_i^2$ (recall that $\theta_i=\mu_i$). The mean and variance become:

$E(Y_i)=b'(\theta_i)=\theta_i=\mu_i$ and

$Var(Y_i)=b''(\theta_i)a_i(\varphi)=\sigma^2$ respectively.

2.0.3 The Link Function

The second element of the generalization is that instead of modeling the mean, as before, we will introduce a one-to-one continuously differentiable (up to order n) transformation g (μ_i) and focus on:

$$\eta_i=g(\mu_i) \qquad (7)$$

The function g (μ_i) will be called the link function. Examples of link functions include the identity, log, reciprocal, logit and probit functions. We further assume that the transformed mean follows a linear model, so that

$$\eta_i=x'\beta \qquad (8)$$

The quantity η_i is called the linear predictor. Note that the model for η_i is pleasantly simple. Since the link function is one-to-one we can invert it to obtain $\mu_i=g^{-1}(x_i'\beta)$. (The model for μ_i is usually more complicated than the model for η_i). We do not transform the response y_i, but rather its expected value μ_i. A model where log y_i is linear on x_i, for example, is not the same as a generalized linear model where log μ_i is linear on x_i. When the link function makes the linear predictor η_i the same as the canonical parameter θ_i, we say that we have a canonical link. The identity is the canonical link for the normal distribution. In the same vein, the logit is the canonical link for the binomial distribution and the log is the canonical link for the Poisson distribution. Other combinations are also possible. An advantage of canonical links is that a minimal sufficient statistic for **β** exists, i.e. all the information about **β** is contained in a function of the data of the same dimensionality. The use of link functions is essential in poverty studies where a response variable is usually binary (0=non-poor, 1=poor) and especially for cases where prediction may run wild i.e. where values are outside the range of (0, 1). The link function thus relates the linear predictor to the expected value of the data and is a one-to-one, continuous, differentiable, invertible and an integratable function. Suppose F (.) is a cumulative distribution function (c.d.f.) of a random variable defined on the real line, and $\pi_i = F\left(y_i\right)$ for $-\infty < y_i < \infty$.

The inverse transformation $y_i = F^{-1}\left(\pi_i\right)$ for $0<\pi<1$ is the link function. Popular choices are the logistic and extreme value distribution. Let Y_i denote a random variable representing a binary response coded zero or one, as usual. We call Y_i the manifest response. Suppose that there is an unobservable continuous variable Y_i^* which can take any value in the real line, and such that Y_i takes the value one if and only if Y_i^* exceeds a certain threshold θ. We call Y_i^* the latent response. Since a positive outcome occurs only when the latent response exceeds the threshold, we can write the probability π_i of a positive outcome as $\pi_i = P(Y_i = 1) = P(Y_i^* > \theta)$. As often with latent variables, the location and scale of Y_i^* are arbitrary. We can add a constant to both Y_i^* and the threshold θ, or multiply both by a constant *C*, without changing the probability of a positive outcome. To identify the model we take the threshold to be zero, and standardize Y_i^* to have standard deviation one (or any fixed value). Suppose now that the outcome depends on a vector of covariates **X**. To model this dependence, we use an ordinary linear model for the latent variable by writing $Y_i^* = X_i\beta + \mu_i$ where β is a vector of coefficients of the covariates and μ_i is the error term, assumed to have a distribution with c.d.f. F(U) not necessarily the normal distribution. Under this model, the probability π_i of observing a positive outcome is:

$$\pi_i = P(Y_i > 0) = P(U_i > -y_i) = 1 - F(-y) \quad (9)$$

where $y_i = X_i\beta$ is the linear predictor. If the distribution of the error term U_i is symmetric about zero so that

$$F(U) = 1 - F(-U)$$

Then $\pi_i = F(y_i)$ (10)

This expression defines a generalized linear model with Bernoulli response link: $y_i = F^{-1}(\pi)$. In the more general case where the distribution of the error term is not necessarily symmetric, we still have a generalized linear model with link $y_i = -F^{-1}(1-\pi_i)$.

2.0.4 Maximum Likelihood Estimation

An important practical feature of generalized linear models is that they can all be fit to data using the same algorithm, a form of iteratively re-weighted least squares. In this section we describe the algorithm. Given a trial estimate $\hat{\beta}$ of the parameters, we calculate the estimated linear predictor $\hat{\eta}_i = x'\hat{\beta}$ and use that to obtain the fitted values $\hat{\mu}_i = g^{-1}(\hat{\eta}_i)$. Using these quantities, we calculate the working dependent variable:

$$z_i = \hat{\eta}_i + (y_i - \hat{\mu}_i)\frac{d\eta_i}{d\mu_i} \quad (11)$$

where the rightmost term is the derivative of the link function evaluated at the trial estimate.

Next we calculate the iterative weights

$$w_i = \frac{p_i}{b''(\theta_i)(\frac{d\eta_i}{d\mu_i})^2} \quad (12)$$

where $b''(\theta_i)$ the second derivative of is $b(\theta_i)$ evaluated at the trial estimate and we have assumed that $a_i(\varphi)$ has the usual form $\frac{\varphi}{p_i}$. This weight is inversely proportional to the variance of the working dependent variable z_i given the current estimates of the parameters, with proportionality factor φ. Finally, we obtain an improved estimate of β regressing the working dependent variable z_i on the predictors x_i using the weights w_i, i.e. we calculate the weighted least-squares estimate

$$\hat{\beta} = (X'WX)^{-1}X'WZ \quad (13)$$

where X is the model matrix, W is a diagonal matrix of weights with entries w_i given by (12) and **z** is a response vector with entries z_i given by (11). The procedure is repeated until successive estimates change by less than a specified small amount. [17] has shown that this algorithm is equivalent to Fisher scoring and leads to maximum likelihood estimates. They consider the general case of $a_i(\varphi)$ and include φ in their expression for the iterative weight. In other words, they use $w_i^* = \varphi w_i$, where w_i is the weight used in the discussion here. The proportionality factor φ cancels out when we calculate the weighted least-squares estimates using (13), so the estimator is exactly the same.

3.0 Results and Discussion

The STATA software was employed for the generalized linear modeling of poverty determinants. The per capita expenditure was selected as the dependent variable.

This was coded into poor (1) and non-poor (0). Households with per capita expenditure less than the poverty line, z were deemed poor and those with per capita expenditure greater than the poverty line were regarded as non-poor households. For the generalized linear modeling, link functions from the binomial family were considered. The reason for this is based on the fact that the response variable is dichotomous (0= non-poor and 1 = poor). The various link functions considered were Logit, Probit, Log-log and Complementary log-log respectively. The performance of each link function was assessed based on the value of its deviance statistic. The generalized model yielding the minimum deviance statistic was adjudged best in all cases. The independent variables selected were: Sex (sex of the household head), Ageyrs (age in years of the household head), Fatheduc (father's education level), Fathwrk (father's work), Motheduc (mother's education level), Mothwrk (mother's work), Hhsize (household size), Occgrp (occupation group of the household head), Edgrp (educational group for highest level attained by the household head) and Lit (literacy of the household head). The results for the generalized linear modeling are shown in Table 1-Table 4. The log-log link function performed best as it gave the minimum value of the deviance statistic. The logit, probit and complementary log-log came second, third and fourth respectively. The log-log showed that all the variables were significant at $\alpha = 5\%$ level of significance. The logit, probit and complementary log-log link functions indicated that all the variables except Motheduc (mother's education level) and Lit (literacy of the household head) were significant. The implication of this result is that generally all the selected variables are associated with the poverty status of the household. That is they are all poverty correlates. However the correlation between poverty status of the household and mother's education level as well as the literacy of the household head is weak

Table 1: Results for Generalized Linear Modelling-1
Family=Bernoulli, Link Function= Logit, AIC = 1.20283, Deviance = 23021.71378
Scale Deviance = 1.20237

Variable	Coefficient	Standard Error	Z	P> |Z|	Remark
Constant	-0.64103	0.16456	-3.9	0.000	S
Sex	-0.25712	0.04766	-5.39	0.000	S
Ageyrs	-0.01028	0.00116	-8.83	0.000	S
Fatheduc	0.04308	0.00578	7.45	0.000	S
Fathwrk	0.01141	0.00163	6.99	0.000	S
Motheduc	0.00865	0.00509	1.7	0.090	NS
Mothwrk	-0.00722	0.00194	-3.73	0.000	S
Hhsize	0.29496	0.00705	41.82	0.000	S
Occgrp	0.05594	0.00846	6.61	0.000	S
Edgrp	-0.24548	0.00995	-24.66	0.000	S
Lit	0.08366	0.04542	1.84	0.065	NS

S= significant at α = 5%, NS= insignificant at α = 5%

Table 2: Results for Generalized Linear Modelling- 2
Family=Bernoulli, Link Function= Probit, AIC = 1.20583, Deviance = 23079.40166
Scale Deviance = 1.2053

Variable	Coefficient	Standard Error	Z	P> |Z|	Remark
Constant	-0.64103	0.09893	-3.51	0.000	S
Sex	-0.25712	0.02888	-5.77	0.000	S
Ageyrs	-0.01028	0.00070	-8.83	0.000	S
Fatheduc	0.04308	0.00346	7.34	0.000	S
Fathwrk	0.01141	0.00097	6.94	0.000	S
Motheduc	0.00865	0.00306	1.81	0.071	NS
Mothwrk	-0.00722	0.00117	-3.67	0.000	S
Hhsize	0.29496	0.00390	43.72	0.000	S
Occgrp	0.05594	0.00505	6.91	0.000	S
Edgrp	-0.24548	0.00595	-24.83	0.000	S
Lit	0.08366	0.02737	1.83	0.068	NS

S= significant at α = 5%, NS= insignificant at α = 5%

Table 3: Results for Generalized Linear Modelling- 3
Family=Bernoulli, Link Function= Log-log, AIC = 1.19710, Deviance = 22912.0.5576
Scale Deviance = 1.19664

Variable	Coefficient	Standard Error	Z	P> |Z|	Remark
Constant	-0.07592	0.10775	-0.70	0.481	NS
Sex	-0.13984	0.03027	-4.62	0.000	S
Ageyrs	-0.00659	0.00077	-8.61	0.000	S
Fatheduc	0.03134	0.00422	7.43	0.000	S
Fathwrk	0.00670	0.00102	6.55	0.000	S
Motheduc	0.00733	0.00367	2.00	0.046	S
Mothwrk	-0.00542	0.00129	-4.21	0.000	S
Hhsize	0.23007	0.00523	43.98	0.000	S
Occgrp	0.02706	0.00541	5.01	0.000	S
Edgrp	-0.16968	0.00687	-24.70	0.000	S
Lit	0.06399	0.03243	1.97	0.048	S

S= significant at α = 5%, NS= insignificant at α = 5%

Table 4: Results for Generalized Linear Modelling- 4
Family=Bernoulli, Link Function= Complementary log-log, AIC = 1.21829, Deviance = 23317.97543, Scale Deviance = 1.21784

Variable	Coefficient	Standard Error	Z	P> |Z|	Remark
Constant	-0.67360	0.11522	-5.85	0.000	S
Sex	-0.24780	0.03577	-6.93	0.000	S
Ageyrs	-0.00710	0.00081	-8.82	0.000	S
Fatheduc	0.02521	0.00361	6.99	0.000	S
Fathwrk	0.00769	0.00114	6.73	0.000	S
Motheduc	0.00506	0.00327	1.55	0.121	NS
Mothwrk	-0.00397	0.00131	-3.02	0.002	S
Hhsize	0.15550	0.00351	44.30	0.000	S
Occgrp	0.05144	0.00592	8.69	0.000	S
Edgrp	-0.15827	0.00654	-24.21	0.000	S
Lit	0.05107	0.02926	1.75	0.081	NS

S= significant at α = 5%, NS= insignificant at α = 5%

4.0 Conclusion

This study has demonstrated the suitability of generalized linear modeling in analyzing poverty in Nigeria. It considered the logit, probit, log-log and complementary log-log link functions with the log-log link performing best. The generalized linear models showed that sex of the household head, age in years of the household head, father's education level, father's work, mother's work, household size, occupation group of the household head, and educational group for highest level attained by the household head are strong correlates of poverty while mother's education level and literacy of the household head are weak correlates of poverty. The findings in this study agree with past studies [14, 18-20]. Hence having fewer numbers of children parents can cater for and ensuring education of children born to families by parents should be encouraged to reduce the burden of poverty.

References

[1]. Omonona, B.T, Udoh, E.J, and Adeniran, A.A (2008). Poverty and its Determinants among Nigerian Farming Households: Evidence from Akinyele Lga of Oyo State, Nigeria. *European Journal of Social Sciences, Volume 6, Number 3, Pp 402-412.*

[2]. World Bank (1995). *Advancing Social Development*. Washington, D.C. International Bank for Reconstruction and Development.

[3]. FOS (1999). *Poverty Profile for Nigeria*, 1980-1996. Federal Office of Statistics, Abuja, Nigeria.

[4]. Glewwe, P and Van der Gaag, J. (1988). Confronting Poverty in Developing Countries. Definitions, Information and Policies. *Living Standard Measurement Study Working Paper No.48* Washington D.C. The World Bank.

[5]. World Bank (1990). *World Development Report*. New York: Oxford
` University Press.
[6]. Anthonio, B.O. (1975). Sectoral Analysis of Rural Poverty in Nigeria. *Proceedings of the Nigerian Economic Society's Annual Conference.*
[7]. Oshuntogun, A. (1975). Poverty as an issue in Rural Development Policy. A Case Study from Western State of Nigeria. *Proceedings of the Nigerian Economic Society's Annual Conference.*
[8]. Adeyokunnu, T.O. (1975). Rural Poverty in Nigeria. A Case Study of Ife vision. *Proceedings of the Nigerian Economic Society's Annual Conference.*
[9]. Ogwumike, F.O. (1991). A Basic Needs - Oriented Approach to the Measurement of Poverty. *Nigerian Journal of Economics and Social Studies* 33(2).
[10]. Ogwumike, F.O. and Ekpeyong, D.B. (1996). Impact of Structural Adjustment Policies on Poverty in Nigeria. Research Report AERC, Nairobi.
[11]. Anyanwu, J.C. (1997). Poverty in Nigeria: Concepts, Measurements and Determinants. Selected Papers from the Nigerian Economic Society's Annual Conference.
[12]. Englama, D. and Bamidele, A. (1997). Measurement Issues in Poverty. Selected Papers from the Nigerian Economic Society's Annual Conference.
[13]. Odusola, A.F. (1997). Poverty in Nigeria: An Eclectic Appraisal. Selected papers from the Nigerian Economic Society's Annual Conference.
[14]. Osowole, O.I**.,** Ugbechie, Rita and Uba, Ezenwanyi (2012). On the identification of Core Determinants of Poverty: A Logistic Regression Approach. *Mathematical Theory and Modeling Vol.2, No.10 Pp 45-52.ISSN 2225-0522(Online).www.iiste.org* .
[15]. Ravallion, M. (1996). Issues in Measuring and Modelling Poverty. *The Economic Journal* 106 *(September) Pp 1328-1343.*
[16]. Nelder, J.A. and Wedderburn, R.W.M.1972. Generalized Linear Models. *Royal Journal Statist. Soc.A.135 (Part 3): 370-384.*
[17]. McCullagh, P and Nelder, J.A.1989. *Generalized Linear Models, Second Edition*. London: Chapman and Hall.
[18]. Datt, G. 1998. Computational Tools for Poverty Measurement and Analysis. FCND Discussion Paper No. 50. Food and Consumption and Nutrition Division, International Policy Research Institute 2033k Street, N.W. Washington, D.C., 20006 USA
[19]. Al-Saleh, J.2000. Poverty Assessment in Palestine: Preliminary Report. *Inter-Stat. 22: 83-106*
[20]. Ajakaiye, D.O. and Adeyeye, V.A.2002. Concepts, Measurements and Causes of Poverty. *CBN Economic and Financial Review. Prod. 39:4*