IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 11, Issue 3 Ver. V (May - Jun. 2015), PP 73-87
www.iosrjournals.org

Multidimensional Analysis of the Determinants of Poverty Indicators in the Lake Victoria Basin(Kenya)

[1]Anthony Ngunyi, [2]Peter. N. Mwita, [2]Romanus O. Odhiambo, [3]Verdiana. G. Masanja

[1,]Dedan Kimathi University of Technology (Kenya)P.O BOX 657-10100, Nyeri
[2,] Jomo Kenyatta University of Agriculture & Technology (Kenya) P.O BOX 62000-00200, Nairobi
[3,] University of Rwanda(Head Quarters) P.O Box 4285, Kigali, Rwanda
[3,]Affiliated to University of Dar es Salaam, P.O.Box 35062 ,Dar es Salaam,Tanzania

Abstract: *The study main objective is to examine the multidimensional aspects of poverty in one Kenya's culturally diverse region of the Lake Victoria basin. The analysis using data collected by IUCEA researchers in 2007 and also the 2009 census on households in Kenya. This study investigates statistical models based on factors that characterize the demographic characteristic of individuals, in determining the predictors of poverty for better policy formulation.. The research findings indicate that poverty measures do overlap to capture a percentage of the sample as poor. The analysis shows that education, gender (being male), marital status, assets (livestock, water sources, and wall materials) and age of the head of the family have statistically positive effects on the likelihood of an individual falling into poverty.*

Keywords: *Poverty, Demography, Augmented, Logistic, Assets*

I. Background Information

According to the World Bank,(2010).[1], "poverty is pronounced deprivation in well-being." This of course begs the questions of what is meant by well-being and of what is the reference point against which to measure deprivation.

The objective of the study was to look at the different factors that influence poverty in a household, and the policy formulation that can be put in place in order to achieve bettering living standard for the members of the household. This study based its results on a multidisplinary aspects on the fact that many studies on poverty in Kenya have been on regressing well known determinants even though other factors may be able to give an informative and simple to interpret facts on poverty levels in the region.

One approach is to think of well-being as the command over commodities in general, so people are better off if they have a greater command over resources. The main focus is on whether households or individuals have enough resources to meet their needs, see, S. Pudney(1999) [2]. Typically, poverty is then measured by comparing individuals' income or consumption with some defined threshold below which they are considered to be poor. This is the most conventional view-poverty is seen largely in monetary terms-and is the starting point for most analysis of poverty.

A second approach to well-being (and hence poverty) is to ask whether people are able to obtain a specific type of consumption: Do they have enough food? Or shelter? Or health care? Or education? As cited in Ravallion and Bidani (1994); Kakwani (1990),[3,4]. In this view the analyst goes beyond the more traditional monetary measures of poverty: Nutritional poverty might be measured by examining whether children are stunted or wasted; and educational poverty might be measured by asking whether people are literate or how much formal schooling they have received, well articulated in Lipton and Ravallion (1995),[5].

Perhaps the broadest approach to well-being is the one articulated by Sen (1999), [6], who argues that well-being comes from a capability to function in society. Thus, poverty arises when people lack key capabilities, and so have inadequate incomes or education, or poor health, or insecurity, or low self-confidence, or a sense of powerlessness, or the absence of rights such as freedom of speech. Viewed in this way, poverty is a multidimensional phenomenon and less amenable to simple solutions. For instance, while higher average incomes will certainly help reduce poverty, these may need to be accompanied by measures to empower the poor, or insure them against risks, or to address specific weaknesses such as inadequate availability of schools or a corrupt health service (Datt and Jolliffe, 2005). [7].

WHO (2000),[8] noted that poverty is related to, but distinct from, inequality and vulnerability. Inequality focuses on the distribution of attributes, such as income or consumption, across the whole population. In the context of poverty analysis, inequality requires examination if one believes that the welfare of individuals depends on their economic position relative to others in society. Vulnerability is defined as the risk of falling into poverty in the future, even if the person is not necessarily poor now; it is often associated with the effects of "shocks" such as a drought, a drop in farm prices, or a financial crisis. Vulnerability is a key dimension of well-being since it affects individuals' behavior in terms of investment, production patterns, and coping strategies, and in terms of the perceptions of their own situations.

According to the last Country Briefs, an estimated 3.8 million people in rural areas are between highly to extremely food insecure. Food and Agriculture Organization (FAO)/ Global Information and Early Warning System on Food and Agriculture (GIEWS) and Famine Early Warning System (FEWSNET) agree that, in the short term, Kenya is a hunger-prone country, while WFP and IFPRI assess the long-term situation as alarming and hunger as moderately high.. There is a long history of periodic shortfalls in food supply in Kenya. Shortfalls occur all over the country or in parts of the country, and sometimes for two years in a row. In times of unfavorable weather, even the provinces normally characterized by a maize surplus (such as the Rift Valley) or marginally self-sufficient provinces (such as Western and Nyanza) may enter a maize deficit situation. In addition, in areas characterized by chronic deficits (such as the Coast and Eastern and North Eastern provinces) the situation becomes acute. In many districts in these areas, emergency relief becomes necessary.

The highest poverty rate was found among people living in households headed by farmers 46 percent (KNBS, 2007a), [11]. By contrast, households headed by someone working in the government are least likely to be poor; in these occupations the poverty rate was 20 percent (1993–94). This would suggest that policies that aim to reduce poverty through enhancing income-generating capabilities should be targeted towards the agricultural sector.

The relationship between poverty and education is particularly important because of the key role played by education in raising economic growth and reducing poverty. The better educated have higher incomes and thus are much less likely to be poor. Kenyans living in households with an uneducated household head are more likely to be poor, with a poverty rate of 47 percent in 2014 national poverty atlas.. With higher levels of education, the likelihood of being poor falls considerably. Raising education attainment is clearly a high priority to improve living standards and reduce poverty.

The relationship between gender and poverty may also indicate another targeting strategy for poverty reduction. In Tanzania, about 35 percent of the population lives in households headed by women. Perhaps surprisingly, the 2007 data show that the poverty rate was slightly lower among female-headed households (48 percent) than among male-headed households (52 percent). In this case, targeting interventions based on the gender of the head of household would not help to distinguish the poor from the non-poor, Mark Schreiner, [13].

II. Literature Review

Poverty is a worldwide concern. Although there is a global concern towards poverty reduction, there is a little agreement on a single definition and measurement of poverty (Kotler et al., 2006; Laderchi et al., 2003), [14, 15]. According to Kotler et al., (2006),[14] and Laderchi et al.(2003),[15], the problem of arriving at one single definition of poverty has been compounded by a number of factors. Poverty affects heterogeneous groups such that the concept of poverty is relative depending on different interest groups and individuals experiencing it (Kotler et al., 2006, Rank, 2004), [14, 16]. The difficulty surrounding the definition and measurement of poverty has often led poverty researchers and policy makers to relate poverty to the concepts of impoverishment, deprivation, the disadvantaged, inequality, the underprivileged and the needy.

Many researchers have authored many articles on the issue of poverty worldwide. The exception being the absolute poverty measures for the developing world by Chen and Ravallion (2007) [1], which serve to provide the latest evidence for an African exceptionalism that dominates the development needs of today.

All developing country regions have shown marked improvement in key indicators of poverty,health, economy, and food, except for sub-Saharan Africa. For poverty, the global number of people living below the extreme poverty line of $1 per day decreased between 1981 and 2004 from 1,470 million to 969 million. The percentage of extremely poor fell from 40% to 18%. However, in sub- Saharan Africa, the numbers almost doubled from 168 million to 298 million, and the percentage stayed almost constant from 42% to 41% , Chen S, Ravallion M (2007) [35].

For health, the life expectancy at birth in sub-Saharan Africa peaked in 1990 at 50 years but has since declined to 46 years, while steadily rising in all developing country regions to an average of 65 years, Jamison D.T, (2006),[36]. Over the period 1960–2000, sub-Saharan Africa's per capita measure of annual economic growth (gross domestic product) was a mere 0.1%, whereas other developing country regions experienced accelerated growth averaging 3.6%, Collier P (2007), [37]. Food production per capita grew by 2.3% per year between 1980 and 2000 in Asia, grew by 0.9% in Latin America, and declined by 0.01% in tropical Africa see, Dasgupta .P et al (2004),[38].

There are basically two approaches in modelling determinants of poverty. The first approach[5] is the employment of consumption expenditure per adult equivalent and regress it against potential explanatory variables (Geda et al, 2001). Using this approach Arneberg and Pederson (2001) report that household characteristics and education are the main factors which affect living standard in Eritrea. However, they treat education as a linear and continuous variable. Moreover they find out that transfer payment from relatives abroad is a significant contributor to the welfare of a society. From their analysis they conclude that education is the most important factor for the way out of poverty. However, their approach suffers from the common problems of consumption as being indicator of welfare and the assumption that consumption of the poor and non poor are both determined by the same process (Okwi, 1999). The second approach is to directly model poverty by employing a discrete choice model.

The practice of discrete choice models in the analysis of determinants of poverty has been popular approach[6] (for instance, Fafack(2002) for Burkian'faso, Kabubuo-Mariara (2002) for Kenya; Amuedo_Dorantes(2004) for Chile; Grootaert(1997) for Cote D'voire; Geda et al (2001) for Kenya; Charlette-Gueard and Mesple-Somps (2001) for Cote d`voire , Goaed and Ghazouani (2001) for Tunisia; Roubaud and Razafindrakoto ,2003). The analysis then proceeds by employing binary logit or probit model to estimate the probability of a household being poor conditional up on some characteristics. In some cases also the households are divided into three categories: absolute poor, poor and non poor and then employ ordered logit or ordered logit model to identify the factors which affect the probability a household being poor conditional up on set of characteristics. In this study we apply the dicrete choice model as discussed by many researchers in kenya but also look at the augmented model proposed by Datt. G and Jolliffe .D. (2005),[34]

Common indices developed by the United Nations Development Programme are the human development index composed of three measures of development (per capita gross domestic product, life expectancy, and literacy) or the human poverty index composed of measures of deprivation in the development indices (child and young adult mortality, illiteracy, and lack of water and sanitation) United Nations Development Programme (2006),[37].In the study of the lake Victoria basin ,we look at the aspects of the asset component as a measure of poverty and articulate the best policy measures that can be taken into consideration to reduce poverty in the area.

Poverty studies in Kenya have focused on a discussion of inequality and welfare based on limited house level data (Arne, 1981; Hazlewood, 1981; House and Killick, 1981) [17,18,19]. One recent comprehensive study on the subject is that of Geda et al. (2001), [20], which deals with measurement, profile and determinants of poverty. The study employs a household welfare function, approximated by household expenditure per adult equivalent. The authors runs two categories of regression, using overall expenditures and food expenditures as dependent variables. In each of the two cases, three equations are estimated which differ by type of dependent variable. These dependent variables are: total household expenditure, total household expenditure gap (the difference between the absolute poverty line and the actual expenditure) and the square of the latter. A similar set of dependent variables is used for food expenditure, with the explanatory variables being identical in all cases.

Geda et al. (2001), [20] , justified their choice of this approach (compared to a logit/probit model) as follows; First, the two approaches (discrete and continuous choice-based regressions) yield basically similar results; also the expenditure as a binary variable has certain inherent weakness. One obvious weakness is that, unlike the logit/probit model, the level of the regression about poverty. Second, the major assumption of the welfare function approach is that consumption expenditure are negatively associated with absolute poverty at all expenditure levels. Thus factors that increase consumption expenditure reduce poverty. However, this basic assumption needs to be taken cautiously. For instance though increasing welfare, raising the level of consumption expenditure of households that are already above the poverty line does not affect the poverty level (for example measures by the headcount ratio). Notwithstanding such weakness, the approach is widely used

Geda et al. (2001),[20] identified the following as important determinants of poverty: unobserved region-specific factors, mean age, size of household, place of residence (rural versus urban), level of schooling, livestock holding and sanitary conditions. The importance of these variables does not change whether the total expenditure, the expenditure gap or the square of the gap is taken as the dependant variable. The only noticeable change is that the sizes of the estimated coefficients are enormously reduced in the expenditure gap and in the square of the expenditure gap specifications. Moreover, except for the minor changes in the relative importance of some of the variable, the pattern of coefficient again fundamentally remains unchanged when the regressions are run with food expenditures as dependant variable.

Another recent study on the determinant of poverty is Oyugi (2000),[21], which is an extension to earlier work by Greer and Thorbecke (1986b,a).[22,23]. The later study used household calorie consumption as the dependant variable and a limited number of household characteristics as explanatory variables. An important aspect of Oyugi's study is that it analyse poverty both at micro (household) and meso (district) level, with the meso level analysis being the innovative component of the study. The explanatory variable (household characteristics) include: holding area livestock unit, the proportion of household members able to read and write, household size, sector of economic activity (agriculture, manufacturing/industrial). The results of the probit analysis show that all variable used are important determinants of poverty in rural areas and at the national level, but that there are important exceptions for urban areas.

In the probit model, however, in the order of importance the key determinants of poverty are: being able to read and write, employment in off-farm activities, being engaged in agriculture, having a side-business in the service sector, source of water and household size. Region of residence appears to be equally important in determining poverty status in the two approaches. Although the two approaches did not employ the same explanatory variables, this comparison points to the possibility of arriving at different policy conclusions from the two approaches Oyugi (2000),[21].

III. Methodology

3.1 Area of Study

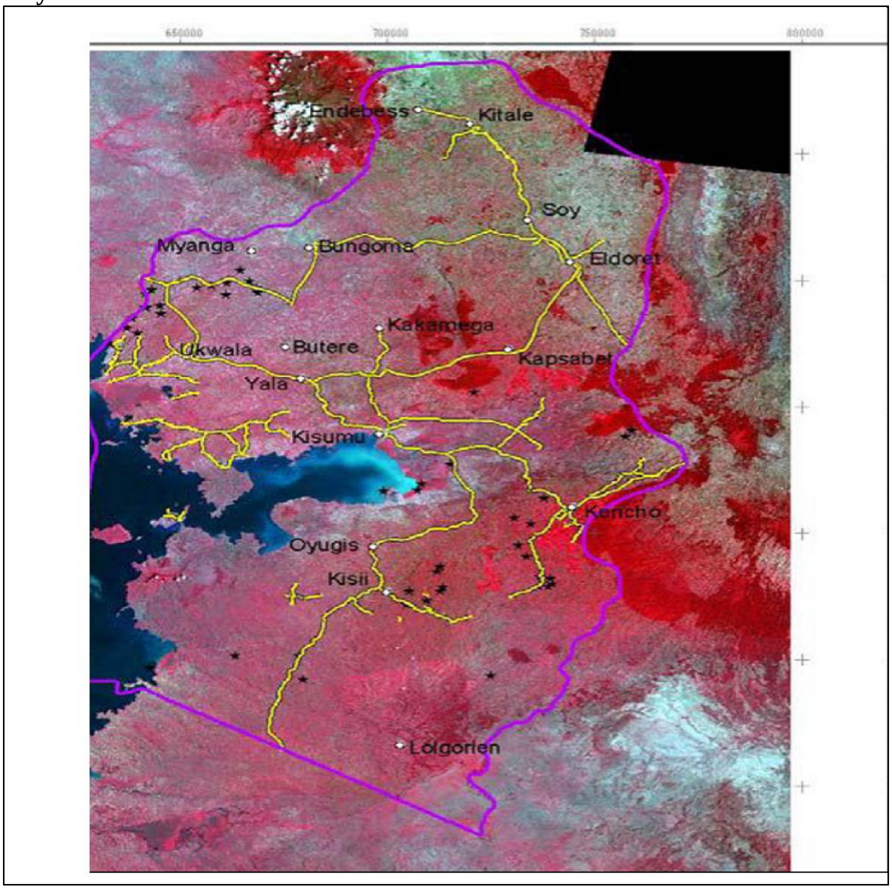

Figure 1: The Lake Victoria Basin on the Kenyan side.

The study site constitutes three districts of the Lake Victoria basin. Some of the raw data on Household Demography, collected during the first year work of the project entitled “Mathematical Techniques for Food Crops Balance Sheet and Food Security Indicators in Lake Victoria Water Shed”, see [24] is used in this study. The random sampling approach was employed to select the study areas and sample respondents in which the subjects selected were supposed to meet the study needs. A total of 24 households in each of the three districts (Kuria, Siaya and Kisumu) in Kenya were surveyed using structured questionnaire, interview sessions, focus group discussion and observation. A list of household heads (which is the sampling frame from which a probability sample is selected) were supplied by respective sub-location administrations. These lists were each used to select 24 households from each sub-location by employing simple random sampling technique. This method of sample selection is free form bias; it has given every household head in each sub-location a chance of being included in the sample for this study.

The study also makes use of data obtained from the 2009 Population and Housing Census conducted by the Kenya National Bureau of Statistics. The survey questionnaire collected information on household and demographic characteristics, education, assets, employment, income, and expenditures and assets in the households. The questionnaire included information on household members and was administered to all households in the country, with the exception of North Eastern Province. Although the census did not collect information on income and expenditures, it provides information on a number of characteristics that have been shown to be strong correlates of poverty. . Such characteristics include assets, education and the household size.

3.2 Specification of the Regression Model

When poverty is defined as the current consumption deficit, a household is categorized as poor if the value of per capita consumption of its members is lower than the poverty line. Therefore, it is logical to search for poverty predictors based on variables that correlate with per capita household consumption. These variables can be obtained by estimating a model of consumption correlates, where the left-hand side is per capita consumption and the right-hand side is a set of variables that is thought of correlating with household consumption. Different from determinants model, in correlates model the endogeneity of the right-hand side variables is not a concern,see Datt and Jolliffe, 2005). [37].

Once the set of the right-hand side variables has been determined, a stepwise regression procedure is employed to estimate the model. The stepwise estimation procedure is used because in the end we want to obtain a manageable number of variables that can be relatively easily collected in practice and at the same time meaningfully used to predict household consumption level and poverty status.

3.3 The Augmented model

The usual approach concerning poverty measurements has historically been to model poverty Directly.The consumption model can be described as the basic model.Futhermore the model of consumption c_j, the determinants of per capita consumption at the household level in the simplest form of a model is as follows

$$log c_j = \boldsymbol{\beta}_\mathbf{j} \mathbf{x}_\mathbf{j} + e_j \tag{1}$$

where x_j is a set of household characteristics and e_j is a random error term.. It has the feature that the marginal effects of the determinants of consumption are constant across households. It is however arguable that there is heterogenity across households and the marginal effects themselves depend on household characteristics. This concern leads us to consider the augmented model that allows for a range of interaction effects and individual specific marginal effects ($\boldsymbol{\beta}_\mathbf{j}$);

$$log c_j = \boldsymbol{\beta}_\mathbf{j} \mathbf{x}_\mathbf{j} + e_j$$

where $\boldsymbol{\beta}_\mathbf{j} = \boldsymbol{\beta}' + \mathbf{x}_\mathbf{j} + \mathbf{e}_\mathbf{j}$ and hence

$$log c_j = \boldsymbol{\beta}' \mathbf{x}_\mathbf{j} + x_j \phi \mathrm{x}_\mathrm{j} + e_j^* \tag{2}$$

This delivers a model with heteroscedastic errors, $e_j^* = e_j + \varepsilon_j$, which is easily allowed for estimating the variance matrix of the model parameters. The model has a generalized quadratic form which is a numerically equivalent second order approximation to any arbitrary twice differentiable function (Fahrmeir and Kaufmann, 1985). [25].

3.4 Specification of the Poverty logistic Model

Choosing an appropriate model and analytical technique depends on the type of variable under investigation. Regression deal with cases where the dependent variable of interest is a continuous variable which we assume, perhaps after an appropriate transformation, to be normally distributed. But in many applications, the dependent variable of interest is not on a continuous scale; it may have only two possible outcomes and therefore can be represented by an indicator variable taking on values 0 and 1.

In this study, the dependent variable Y was defined to have two possible outcomes:

1. The household is poor (1)
2. The household is not poor (0)

These two outcomes are coded 1 and 0 respectively. This shows that the dependent variable is dichotomous and it can be represented by a variable taking the value 1 with probability π and the value 0 with probability $1-\pi$. Such a variable is a point binomial variable, that is, a binomial variable with n =1 trial, and the model often used to express the probability π as a function of potential independent variables under investigation is the logistic regression model. Therefore, to sort out which explanatory variables are most closely related to the dependent variable, nine factors are considered. This method involves a linear combination of the explanatory or independent variables. Thus, the study is modeled within the framework of above mentioned theories and the model used by this study to determine factors affecting poverty status is given equation (3).

3.5 Logistic Regression Analysis

The function has been discussed by many researchers like [26]. It is given by;

$$f(g) = \frac{exp(g)}{1+exp(g)} = \frac{1}{1+exp(g)} \tag{3}$$

when modeling a Bernoulli random variable with multivariate, one directly models the probabilities of group membership, as follows;

$$P(Y=1 \mid X=x) = \frac{1}{1+exp\left(-\left(\beta_0 + \sum_{j=1}^{d} x_j \beta_j\right)\right)} \tag{4}$$

where g in Equation 3 is given by

$$g(X;\beta) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_d X_d \tag{5}$$

To illustrate, the applicability of the logistic function, the bold curve in the figure 2 shows that the logistic function puts more weight on the tails than the normal distribution.

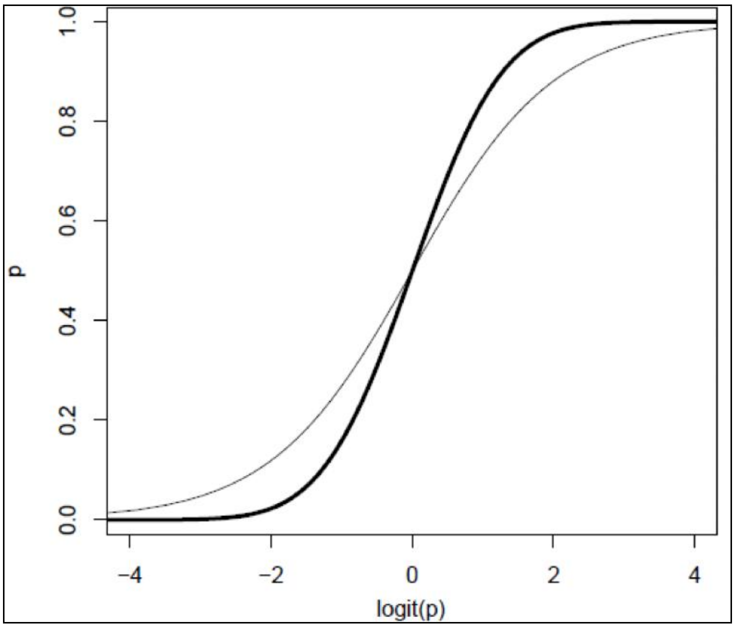

Figure 2: Standardized Normal and Logistic CDF's

Author (2014)

The logistic model is bounded between zero and one, this property estimates the possibility of getting estimated or predicted probabilities outside this range which would not make sense. Also with a proper transformation, one can get a linear model from the logistic function. [26] uses the logit function of the Bernoulli distributed response variable. Transforming Equation 4 as in [26] we have ;

$$Logit[P(Y=1 \mid X=x)] = log_e \frac{P(Y=1 \mid X=x)}{1-P(Y=1 \mid X=x)}$$

$$= log_e \left\{ \frac{1+exp\left(\beta_0 + \sum_{j=1}^{d}\beta_j X_j\right)}{1+exp\left(-\left(\beta_0 + \sum_{j=1}^{d}\beta_j X_j\right)\right)} \right\}$$

$$= log_e \left(exp\left(\beta_0 + \sum_{j=1}^{d}\beta_j X_j\right)\right)$$

$$= \beta_0 + \sum_{j=1}^{d}\beta_j X_j \qquad (6)$$

the function in Equation 6 is a generalized linear model (GLM) with d independent variables.

The motivation to the use of logistic model is that it follows the properties of the GLM. Lets define the hypothetical population proportion of cells for which $Y = 1$ as $\pi = P(Y = 1 \mid X = x)$. Then the theoretical proportion of cells for which $Y = 0$ is $1 - \pi = P(Y = 0 \mid X = x)$. We estimate π by the sample proportions of cells for which $Y = 1$. In the GLM context, it is assumed that there exists a set of predictor variables, $X_1, X_2, \cdots, X_d$, that are related to Y and therefore provides additional information for estimating Y. For mathematical reasons of additivity and multiplicity, logistic model is based on linear model for the log odds in favour of $Y = 1$.

$$log_e \frac{\pi_i}{1-\pi_i} = \alpha + \sum_{j=1}^{d}\beta_j X_j \qquad (7)$$

thus

$$\pi_i = \sum_{j=0}^{d}\beta_j X_j \qquad (8)$$

where $\beta \in \Re^d$ of unknown parameters.

The logistic regression (logit link),

$$g(\pi_i) = log_e \frac{\pi_i}{1-\pi_i} = logit(\pi_i)$$

and

$$g^{-1}(g(\pi_i)) = \pi_i$$

thus the inverse of the logit function in terms of $(X;\beta)$ is given by;

$$g^{-1}(X;\beta) = \pi_i = \frac{1}{1+exp\left(-\left(\beta_0 + \sum_{j=1}^{d}x_j\beta_j\right)\right)}$$

This model can be rewritten as

$$logit(\pi_i) = \sum_{j=0}^{d}\beta_j X_j \qquad (9)$$

IV. Results and Discussion

4.1 Empirical studies of the Stepwise Regression Model and the Augmented Regression Model

For several of the explanatory variables, there are observations with missing data and have constructed dummy variable that take a value of one if the household is missing data for a particular variable(while the value of that variable itself is set as zero). In this way, we reduce the potential of sample selection bias, and we do not miss out on useful information from household with some valid data for most variables.

Per capita consumption is used as the basic measure of individual welfare. The use of per capita consumption imposes the assumptions that there are no economies of household size in consumption and that household composition does not matter, and therefore, the estimated parameters must be intepreted with caution.

There may also be some concern of potential bias in parameter estimates due to endogeneity of omitted variables. If these factors are significant determinants of welfare, the error term will not converge to zero in probability limit and the parameter estimated for the individual explanatory variables will be inconsistent. To control this, interactions term effects are included in the model.

While the augmented equation 2 offers a fairly general approach to modeling welfare, this generality comes with the potential cost of overparameterizing the model with the full set of interaction terms, there are an explosion of parameters. Beginning with a k-parameters in the basic model, there are $\frac{2k+k(k-1)}{2}$ parameters in the augmented equation 2 .

A model with numerous parameters is likely to suffer from multicollinearity. In the view of these difficulties; we use the stepwise regression as our basic model so as to limit them to only those significant in the model. see Micheal. H.K et al.(2005),[39]

Table 2: Stepwise and augmented modeling of the log per capita consumption

Variables	Description	Stepwise model		Augmented model	
		Coefficient	t-ratio	Coefficient	t-ratio
X1	Hhsize	0.4079(.)	1.963	2.297(***)	5.466
X2	Hh size ^{2}	-0.028(*)	-2.062	-0.2410(**)	-3.356
X3	Gender Hh (head)	0.4988(.)	1.853		
X5	Land size (acre)	0.5824(.)	1.983	0.5335(*)	2.203
X6	Hh (head)age	0,1588(***)	3.575		
X7	Hh (head) age ^{2}	-0.0016(**)	-3.575		
X8	Hh Aveage in school	0.0857(.)	1.868	0.2568(**)	2.886
X9	Production(kg) per year	0.0005(***)	1.890	0.0029(*)	2.676
X1:X8	Hh size* Hh Aveage in school			-0.0277(*)	-2.139

(Significance codes: *** 0.001 , **0.01 , *0.05 . 0.01)

(Hh-Household: A domestic unit consisting of members of a family who live together)

Table 2 represents both the stepwise regression model and the augmented model. The null hypothesis, that interactions in the augmented model are jointly equal to zero is convincingly rejected. Thus, there is no support for the standards are uniform across households.

The household size has significant negative (though nonlinear) effects on welfare. This inverse relation between household size and the log per capita consumption is a common finding in the literature (Lanjouw and Ravallion, 1995; Lipton, 2001), [27,28]. The measure of per capita consumption as used in the study is the total food consumption ,non-food and othe expenses of the household. Each of these components of consumption is well documented in more details in the basic report of well-being in Kenya 2005/06 .thus consumption is critically dependent on the underlying assumption regarding economies of household size and equivalent scales.

Education variable emerge as a strong determinant of welfare. In both models the average years of schooling specified on its own have significant positive effects on per capita consumption. However, once the models have been augmented with interactions, several interaction terms in schooling are found to be significant. For example, the marginal return to school is found to be increasing with household size as well as decreasing with the number of the years in school.

We find a strong positive significance effects on the average number of years in education for the family. The models indicate strong positive effects on household if the family is educated. Oduro et al. (2004) ,[29] argue that education and skill acquisition are critical factors for explaining the pattern of rural poverty. Education contributes to the process of moulding attitudinal skills and developing technical skills, and also facilitates the adoption and modification of technology [29].

The study finds that family that owned land (for production) has a significant positive effect on per capita consumption of the household, The age of the household head shows that the expected life cycle in the stepwise model increases poverty status by 15%, also the quadratic term of the age which is nonlinear shows a decline in the life cycle phenomenon of high earning capacity with greater experience and smoothing of consumption over life cycle. There have been similar finding by other authors though using a different techniques, (Datt and Jolliffe, 2005; Mwabu et al., 2000; Oyugi, 2000), [7,30,31].

Table 3: Suitability of the models as indicator of poverty

Models	R^2	Standard error
Stepwise	0.9917	0.815
Augmented	0.9946	0.6895

4.2 Empirical studies of the Logistic Model

This method predicts poverty directly because of the nature of the dependent variable. There are two things that need to be reiterated. First, the dependent variable takes values the values of 1 when the respondent is poor and 0 otherwise.This means in interpreting the estimation result it is important to remember that a positive coefficient means that the variable is correlated positively with the poor. Second, predicted value of the dependent variable is the probability of the observation to be poor.

A logit model has been estimated to elicit the factors influencing welfare status of households. The model uses current welfare status of household as the dichotomous dependent variable. poverty variable is defined on the basis of the variable determinant of poverty indicated below.

The variables in this case are:

Y_i	Poverty of household i ($1=$ Poor, and $0=$ Non-Poor)
X_1	Household size
X_2	Square of household size
X_3	Gender of household head ($1=$ male, and $0=$ female)
X_4	land size(acres)
X_5	Education of HH head ($1=$Primary level and above, $0=$No Education)
X_6	Age of Hh (head)
X_7	Square of Age of Hh (head)
X_8	Per capita aggregate production (No. of Kgs)

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 stepwise method , and the result showed that

The optimal model

$$\begin{aligned} Z_R &= -1.4721\ X_1 + 0.1398\ X_2 + 1.6905\ X_3 + 0.0358\ X_4 \\ &+ 0.0781\ X_5 + 0.3796\ X_6 - 0.0059\ X_7 - 0.3659\ X_8 \end{aligned} \tag{10}$$

According to the model, equation 10, the log of the odds of a household being poor was negatively related to size of the household p=0.01), which according to literature, Paddy (2003) [31] noted that household size was negatively correlated to poverty and Deaton and Paxson (1995) [32] found that food requirement increased in relation to the number of persons in household. The non-linear component of the household size is positively correlated to poverty. This is a common finding in the literature, see [27] and [28].

The log of odds of the gender of the head of the household was positively related to the poverty $(p = 0.05)$.

The age of the household head shows the expected life expectancy. In our model, household living standard increases with the age of the household head upto the optimal age of around 60 years but decreases with the quadratic term which is significant $p = 0.05$. This is consistent with higher earning with greater experience.

There is a strong intergenerational effect on education. Parental education has a strong positive correlation on household welfare.

Food production was expected to be increased extensively through expansion of areas under utilization. The model indicates land size increased food security with 0.4290 even though $(p > 0.05)$. The model, the log odds of land size in positively related to poverty $(p = 0.05)$. In other words, the larger the size of land the increase to production. The production (kg) of the household, the log of the odds indicates that a unit increase of food production improved the food poverty status of the household by 1.4%, with $(p > 0.05)$.

Table 4: Predictors

Predictors	β	$SE(\beta)$	z	p **-value**	e^{β} **(Odds ratios)**
Size of Hh (numbers)	-1.4721	0.0907	-16.230	0.1090	0.2294
Square of household size	0.1398	0.0810	1.7235	0.0844	1.1500
Gender of Hh head (1-male, 0-female)	1.6905	0.8790	1.9232	0.0545	1.0560
Land size(acres)	0.0358	0.2449	0.1461	0.8836	2.4196
Education of Hh head ($1 =$ Primary level and above, $0 =$ No Education)	0.0781	0.1054	0.7419	0.4587	1.5820
Age of of Hh head	0.3796	0.1768	2.1670	0.0318 *	1.0000
Square of Age of Hh head	-0.0059	0.0026	-2.2192	0.0257 *	1.0260
Per capita aggregate production (No. of Kgs)	0.3659	0.1671	2.1897	0.0139^{*}	1.0140

(Dispersion parameter for binomial family taken to be 1)

4.3 Evaluation of the logistic regression model

The overall model evaluation is said to provide a better fit to the data if it demonstrates an improvement over the intercept only model (also called the null model). An intercept only model serves as a good baseline because it contains no predictors. According to this model, all observations would be predicted to belong in the largest outcome category. An improvement over this baseline is examined by using three inferential statistical tests.

Table 5: Statistical inference table Statistical test

Statistics Test	χ^2	df	p
Likelihood ratio test	9.0353	5	0.0854
Hosmer-Lemeshow	9.6702	5	0.7418
Wald test	4.0456	5	0.5443

The statistical significance of individual regression coefficient i.e. (β 's) is tested using the Wald chi-square statistic. According to table 5, the variables are significant predictors of poverty $(p < 0.05)$.

Goodness-of-fit statistics assess the fit of a model against actual values. The inferential goodness-of-fit test is the Hosmer-Lemeshow (H-L) test that yields a $\chi^2_{(5)}$ of 9.6702 and was insignificant $(p < 0.05)$. Suggesting that the model fits the data well. In other word's, the null hypothesis model of a good model fit to data was tenable. The likelihood ratio test yields a $\chi^2_{(5)}$ of 9.0353 and was significant at $p > 0.05$ which also give a good fit for the model and thus the null hypothesis was also tenable for the model.

Table 6: 95% confidence interval for one unit change in X_t

Size of Hh (Number)	$-3.7940, 0.0835$
Square of household size	$0.0002, 0.3473$
Gender of Hh head (1-Male, 0-Female)	$0.1099, 3.8349$
Land size	$-0.4561, 0.3260$
Education of Hh head ($1 =$ Primary level and above, $0 =$ No Education	$-1.0770, 2.5661$
Age of Hh head	$0.0930, 0.8789$
Square of Age of Hh head	$-0.0120, 0.0014$
Per capita aggregate production (kg)	$-0.7457, -0.1033$

The full model is:

$$Z_F = -2.5237\ X_1 + 0.2237\ X_2 + 2.4270\ X_3 + 0.0358\ X_4$$
$$+ 0.7810\ X_5 + 0.6391\ X_6 - 0.00932\ X_7 - 0.3834\ X_8 - 2.076\ X_{11} - 0.0024\ X_{13} \qquad (11)$$

We wish to test

$$H_0 : \beta_0 = \beta_1 = \beta_2 = \cdots \beta_{10}$$
$$H_A : \beta_j \neq 0$$

The reduced model is:

$$Z_R = -1.4721\ X_1 + 0.1398\ X_2 + 1.6905\ X_3 + 0.0358\ X_4$$
$$+ 0.0781\ X_5 + 0.3796\ X_6 - 0.0059\ X_7 - 0.3659\ X_8$$

Table 7: Deviance analysis of the model

Model	Null Deviance	df	Residual Deviance	df
Full model	66.542	48	40.373	35
Reduced model	44.317	32	24.405	24

Therefore, we do not reject the hypothesis, and conclude that the reduced model is a better model than the full model.

4.4 Comparison of the two models using the confusion matrix

The confusion matrix is commonly used to compare two models on how good the predicted respondents. In our study the following matrix were obtained:

Table 8: Logistic model

Indicator observed		
	1	**0**
1	35	0
0	0	23

Table 9: Augmented model

Indicator observed		
	1	**0**
1	34	0
0	0	23

The confusion matrix informs us that the logistic model is better for predicting poverty than the augmented model since it has a high prediction of accurate respondents than the augmented.

4.5 Housing conditions

4.5.1 Roofing Material as measure of poverty

Majority of the respondent represented by 78% stay in corrugated iron sheet houses, followed by with glass thatched houses at 26%, there are also about 1% houses roofed with tiles, another 2.5% with asbestos and the other with about 3% roofed by other materials, this factor may not give a good indicator of poverty but if looked from the perspective of the whole house building material we will be able to see that this indicator can be able to give some indication of poverty.

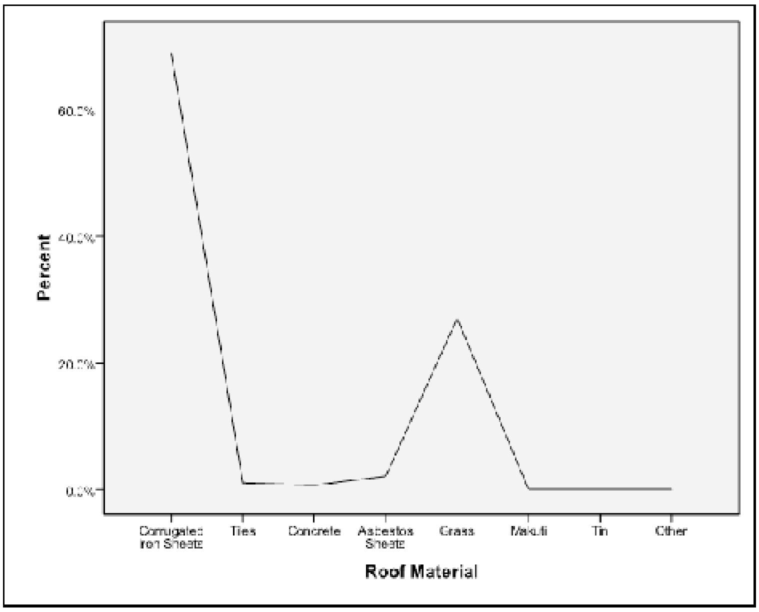

Figure 3: Roofing Materials

4.5.2 Wall material as a measure of poverty

The Majority of houses are walled using mud and wood which represents 62%, 19% are made of bricks, 17% are walled with mud and cement and the others about 3% are walled with other materials like timber and stone which indicates that even combined with roofing materials this area poverty is very high.

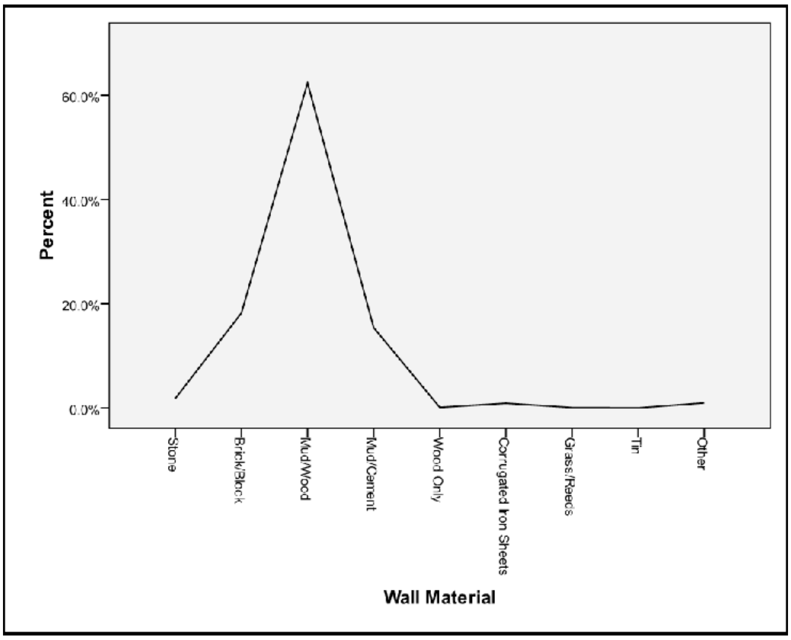

Figure 4: Wall Materials

4.5.3 Main water sources as a measure of poverty

According to [8] about 1.1 billion people lack access to improved water sources, which represents 17% of the global population. In order to achieve the millenium goals, many efforts needs to be done in the areas to ensure the people have clean and safe water. The area majority about 80% only get water from rivers, lake and streams which many times are not clean. [33] also argues that limited access to basic services such as to running water, sanitation on site, grid electricity and health care services is an impediment to escaping from poverty.

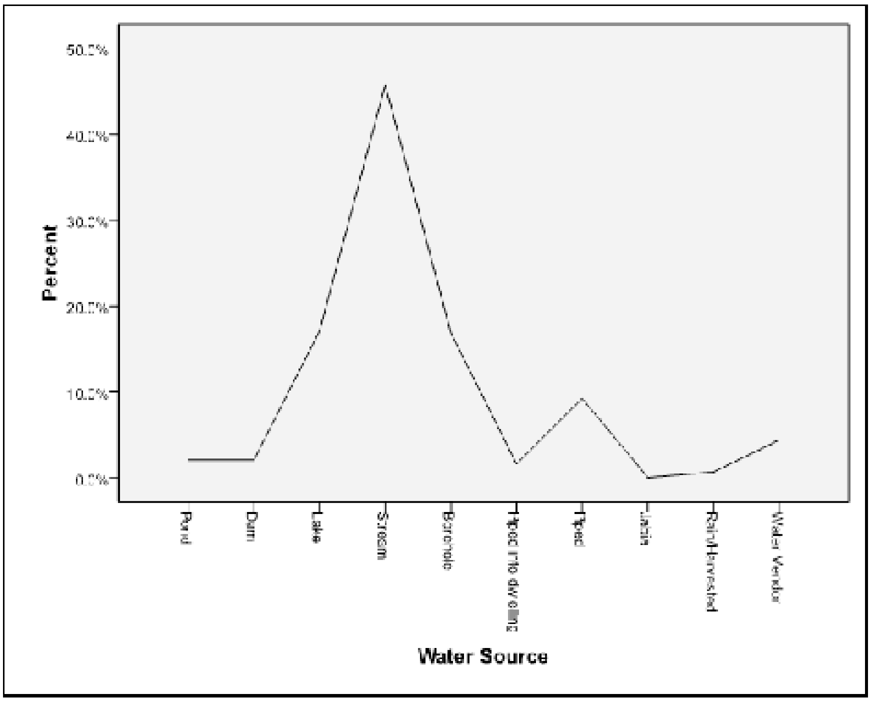

Figure 5: Water Sources

4.6 Information regarding livestock

4.6.1 Poverty against Indigenous cattle

Table 8: Indegenous cattle Chi-square Test

Tests	Value	d.f	Asy. Significance
Pearson chi-square	155.835 [a]	13	0.000
Likelihood ratio test	152.524	13	0.000
Linear by Linear association	73.569	1	0.000
N of valid cases	4414		

[a] 0 cells (0%) have expected count less than 5. The minimum expected count is 7.51

In the table 7, we can see that $chi-squared\ test\ (13)=155.835$ at $p<0.05$. Since the p-value is less than 0.05, we reject the null hypothesis and say that there is statistically significant association between poverty and the rearing of the indigenous cattle in the region. The sample size requirement for chi-squared test of independence is satisfied since zero cells (0 %) has expected count less than 5.

4.5.2 Poverty against Goat

Table 9: Goat Chi-square Test

Tests	Value	d.f	Asy. significance
Pearson chi-square	85.213 [a]	11	0.000
Likelihood ratio test	82.091	11	0.000
Linear by Linear association	45.524	1	0.000
N of valid cases	4414		

[a] 0 cells (0%) have expected count less than 5. The minimum expected count is 5.23

The table 8, shows the relationship between poverty and goat rearing is also statistically significant as we can see from the $chi - squared\ test(11) = 85.213$ at $p < 0.05$.The sample size requirement for chi-squared test of independence is satisfied since zero cells (0 %) has expected count less than 5.

4.5.3 Poverty against Sheep

Table 10: Sheep Chi-square Test

Tests	Value	d.f	Asy. significance
Pearson chi-square	30.444 [a]	8	0.000
Likelihood ratio test	29.185	8	0.000
Linear by Linear association	16.543	1	0.000
N of valid cases	4414		

[a] 0 cells (0%) have expected count less than 5. The minimum expected count is 9.80

Table 9 indicates also that in the region there exists a relationship between poverty and sheep rearing which is statistically significant with $chi - squared\ test\ (8) = 30.444$ at $p < 0.05$.The sample size requirement for chi-squared test of independence is satisfied since zero cells (0 %) has expected count less than 5.

Number of total livestock units owned reduce household poverty rates, implying that assets are important determinants of poverty. This finding is consistent with earlier findings for Kenya [20, 21 and 30].

V. Conclusion

The main objective of the study even with difficulties of obtaining expenditure and income data household precise data and to finding variables that predict poverty in rural areas of Kenya is achieved. In the study we explore the two methods, augmented regression model and the logistic regression model, on predicting poverty .The logistic model was better since it was able to predict correctly all respondents, while the augmented model had a prediction rate of about 2% of not predicting correctly the respondents in the consumption model. However, since our aim is to predict the poor for policy mitigation we focus on the method that provides us with the most accurate prediction. In predicting the poor the logistic model is the best of the consumption models.

Further, we also notice that the variables with the strongest either positive or negative are Land, education, size of the household, age of household head and gender. Furthermore, house characteristics, access to facility and assets play significant role. Thus, if we want to roughly assess whether a household is more likely to be poor or not in the region, it would be better to gather information on assets ownership, education level and consumption patterns as they are the best indicators that should be used to tell the status of poverty in a household. Considering the current population growth rate of about 2.5 percent per annum, there is need for a general overview of the policies to boost economic growth and measures to ensure reduction of poverty to the majority of Kenyans. This should be combined with promotion of family planning to ensure that economic gains and reduced burden on households, as a result of free or subsidized services (e.g. in education and health), do not translate to higher population growth. There is also need for targeted investments in infrastructure such as roads, rural electrification, safety net programmes and provision of water, especially in the marginal areas. The policies on poverty levels in the lake region under the PRSP's three pillar strategy of raising the income opportunities for the poor should focus mostly on agriculture, since the macroeconomic environment is important in determining the productivity which is key to poverty reduction.

References

[1]. World Bank (WB), "World development indicators 2010," Technical report License: CC BY 3.0 IGO, World Bank, Washington, DC, 2010. Available at: https://openknowledge.worldbank.org/handle/10986/4373.
[2]. S. Pudney, "On some statistical methods for modelling the incidence of poverty," Oxford Bulletin of Economics and Statistics, vol. 61, no. 3, pp. 385–408, 1999.
[3]. M. Ravallion and B. Bidani, "How robust is a poverty profile?," World Bank Economic Review, vol. 8, no. 1, pp. 75–102., 1994.
[4]. N. Kakwani, "Poverty and economic growth with an application to cote d'lvoire," LSMS Working Paper 63, World Bank, Washington, D.C, 1990.
[5]. M. Lipton and M. Ravallion, Handbook of Development Economic, vol. 3 of B. Amsterdam: Elsevier, 1995.
[6]. A. Sen, Development as Freedom. Oxford: Oxford University Press, 1999.
[7]. G. Datt and D. Jolliffe, "Poverty in egypt: Modeling and policy simulations," Economic Development and Cultural Change, vol. 53, pp. 327–346, January 2005 2005.
[8]. WHO, "Energy and protein requirements," Tech. Rep. 724, World Health Organization, Geneva, 2000.
[9]. M. Ravallion, ""poverty lines in theory and practice",living standards measurements surveys," tech. rep., World Bank, 1998.
[10]. A. M. P. Prescott, Nicholas, "a poverty profile of cambodia," Discussion Paper 373, World Bank, Washington, DC, 1997.

[11]. KNBS, "Basic report on well being in kenya based on kenya integrated household budgets survey, 2005/2006," tech. rep., Kenya National Bureau of Statistics, 2007.
[12]. KNBS, "Geographic dimension of well being in kenya," tech. rep., Kenya National Bureau of Statistics, 2007.
[13]. Schreiner, Mark. "A Simple Poverty Scorecard for Tanzania", http://www.microfinance.com/English/Papers/Scoring_Poverty_Tanzania EN_2013.pdf, retrieved 31 January 2015..
[14]. P. Kotler, N. Roberto, and T. Leisner, "Alleviating poverty: A macro/micro marketing perspective," Journal of Macromarketing, vol. 26, no. 3, pp. 233–238, 2006.
[15]. C. Laderchi, R. Saith, and F. Stewart, "Does it matter that we do not agree on the definition of poverty: A comparison of four approaches.," Oxford Development Studies, vol. 31, no. 3, pp. 233–274, 2003.
[16]. M. Rank, One Nation Underprivileged: Why American Poverty Affects Us All. New York, NY: Oxford Press., 2004.
[17]. B. Arne, Regional inequality in Kenya. Heinemann Eduacational Books Ltd., 1981.
[18]. A. Hazlewood, 'Income Distribution and Poverty;An Unfortunate View' in Tony Killick. Heinemann Eduacational Books Ltd, 1981.
[19]. W. House and T. Killick, 'Inequality and Poverty in Rural Economy and the Influence of some Aspects of Policy'. Heinemann Eduacational Books Ltd, 1981.
[20]. A. Geda, N. de Jong, G. Mwabu, and M. S. Kimenyi, "Determinants of poverty inkenya: Household-level analysis," Discussion Paper Series 9, The Kenya Institute for Public Policy Research and Analysis (KIPPRA), Nairobi, 2001.
[21]. L. N. Oyugi, "The determinants of poverty in kenya," Master's thesis, Department of Economics , University of Nairobi, 2000. MA Thesis, Department of Economics ,University of Nairobi.
[22]. J. Greer and E. Thorbecke, "A methodology for measuring food poverty applied in kenya," Journal of Developement Economics, vol. Vol. 24, pp. pp. 59–74, 1986.
[23]. J. Greer and E. Thorbecke, "Food poverty profile applied in kenya smallholders," Economic Development and Cultural Change, vol. 35, no. 1, pp. 115–141, 1986.
[24]. P. Mwita, V. Masanja, C. Muyanja, and R. Odhiambo, "Mathematical techniques for food crops balance sheet and food security indicators in lake victoria watershed," tech. rep., Inter-University Council of East Africa, Arusha, 2007.
[25]. I. Fahrmeir and H. Kaufmann, "Consistency and asymptotic normality of maximum likelihood estimator in generalized linear models," Annals of Statistics, vol. 13, pp. 342–368, 1985.
[26]. J. Fan, M. Farmen, and J. Gijbels, "Local maximum likelihood estimator and inference," Stat. Soc. Ser. B Stat. Methodol., vol. 60, no. 3, pp. 591–608, 1998.
[27]. P. Lanjouw and M. Ravallion, "Poverty and household size," The Economic Journal, vol. 105, no. 433, pp. 1415–1434, 1995.
[28]. M. Lipton, "Challenges to meet: food and nutrition security in the new millennium," Nutrition Society, vol. 60, no. 2, pp. 203–214, 2001.
[29]. A. D. Oduro, I. Osei-Akoto, and I. Acquaye, "Poverty in a globalizing economy: The role of rural institutions," tech. rep., FASID, Japan, 2004.
[30]. G. Mwabu, W. Masai, R. Gesami, J. Kirimi, F. Munene, M. Chemengich, and J. Mariara, "Poverty in kenya; profiles and determinants," tech. rep., University of Nairobi and Ministry of Finance & Planning, Mimeo, Nairobi, 2000.
[31]. F. Paddy, "Gender differentials in land ownership and their impact on household food security: A case study of masaka," Master's thesis, Uni–Hohenheim, Uganda, 2003. http:www.troz.uni–hohenheim.de/research/thesis/MSCAES/paddy.pdf.
[32]. A. Deaton and C. Paxson, "Measuring poverty among the elderly," tech. rep., National Bureau of Economic Research, 1995.
[33]. S. Van der Berg, "Poverty and the role of rural institutions in a globalizing south african economy," tech. rep., FASID, Japan, 2004.
[34]. Datt G & Jolliffe D , Poverty in Egypt: Modeling and Policy Simulations, Economic Development and Cultural Change, Vol. 53, pp. 327–346, January 2005
[35]. Ravallion, M., Chen, S., and Sangraula, P. (2007). New evidence on the urbanization of global poverty. Population and Development Review, 33(4):667–701.
[36]. Collier, P. (2007). The bottom billion: Why the poorest countries are failing and what can be done about it. Oxford University Press.
[37]. Jamison, D. T., Breman, J. G., Measham, A. R., Alleyne, G., Claeson, M., Evans, D. B., Jha, P., Mills, A., and Musgrove, P. (2006). Disease control priorities in developing countries. World Bank Publications.
[38]. Arrow, K., Dasgupta, P., Goulder, L., Daily, G., Ehrlich, P., Heal, G., Levin, S., Mäler, K.-G., Schneider, S., Starrett, D., et al. (2004). Are we consuming too much? Journal of Economic Perspectives, pages 147–172.
[39]. Micheal, K. H. and Li, W. (2005). Applied Linear Statistical Models. McGraw-Hill, Irwin, 5th edition.