64

Addis Ababa University

Faculty of Natural Science

Department of Statistics

Household Level Determinants of Poverty in Ethiopia

By: Teshome Kebede Deressa

A Thesis Submitted to the Department of Statistics

Presented in Partial Fullfilment of the Requirements for the

Degree of Master of Science in Statistics

Addis Ababa University

Addis Ababa, Ethiopia

June 2013

Addis Ababa University

Faculty of Natural Science

Department of Statistics

This is to certify that the thesis prepared by Teshome Kebede, entitled: *Household Level Determinants of Poverty in Ethiopia: A Logistic Regression Analysis of HICES 2010-11 Data Set* and submitted in partial fulfillment of the requirements for the Degree of Master of Science in Statistics complies with the regulations of the university and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Advisor: M.K. Sharma (Professor) Signature: M K Sh Date: 27/06/13

Examiner: Eshetu Wencheko (Professor) Signature: ______ Date: 26/06/2013

Examiner: Mekonen Tadesse (Ass. Prof.) Signature: ______ Date: 26/06/2013

ABSTRACT

Household Level Determinants of Poverty in Ethiopia

Teshome Kebede

Addis Ababa University, 2013

Poverty has turned out to be a great global social and economic problem. In Ethiopia, it is multifaceted and deep rooted. The objective of this study was to analyze the impact of socio-economic and demographic characteristics of households on poverty in Ethiopia, both in urban and rural areas, using the latest Household Income, Consumption and Expenditure Survey (HICES) 2010-11 data. The study used a logistic regression model to identify determinants of wellbeing of the household by considering poverty status as dependent variable. Different households are classified as either poor or not-poor on the basis of absolute poverty line yearly per capita consumption of Birr 3781. Results showed that households that own agricultural land, headed by educated person and head of the household employed (self-employed or employed in the formal sector) are more likely to exit from poverty trap. The results of the logistic regression also implied that household wellbeing is negatively affected by being headed by female, having large household size and high dependency ratio. Moreover, the study revealed that the probability of being poor is higher in rural areas. Any policy or program that is geared towards eradicating poverty in the country must recognize the impact of these factors in order for it to succeed.

ACKNOWLEDGMENT

I would never have reached the final destination without the help and support of a number of people whom I have the pleasure and honor to acknowledge in these few paragraphs.

My warmest appreciation goes to my advisor, Professor M.K Sharma, for his constructive support throughout this study. I warmly thank him for his valuable advice and feedback. The work under his guidance gave me the opportunity to understand the area of socio-economic inequalities. The many discussions I have had with him on the topic have helped me to shape my research question.

My thanks also go to all the employees of CSA, especially Ato Abera Eshete for making the data available, for all the support and contributions. My appreciation also extends to my employer, Haramaya University, for the sponsorship that allowed me to pursue my postgraduate education at Addis Ababa University.

I would also like to take this opportunity to thank a number of people who have played an important role during my coursework: Gizachew Kebede, Geleta Abdisa, Tadele Dheressa, Abdeta Gutu and Gutu Adugna. Their energy, their enthusiasm and their friendship helped me to navigate through this long process.

Last but certainly not the least, my immense gratitude goes to my beloved wife, Aberash Getachew, whose unconditional love and support have kept my batteries charged. Her patience and her insights have been priceless.

CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
AUC	Area Under the ROC Curve
CBN	Cost of Basic Needs
CSA	Central Statistical Agency
EA	Enumeration Areas
EDHS	Ethiopia Demographic and Health Survey
ERHS	Ethiopian Rural Household Survey
EUHS	Ethiopian Urban Household Survey
FEI	Food Energy Intake
HDI	Human Development Index
HICES	Household Income, Consumption and Expenditure Survey
H-L	Hosmer and Lemeshow
LL	Log-Likelihood
LRT	Likelihood Ratio Test
MDG	Millennium Development Goals
MoFED	Ministry of Finance and Economic Development
PASDEP	Plan for Accelerated and Sustained Development to End Poverty
PRSP	Poverty Reduction Strategy Paper
ROC	Receiver Operating Characteristics
SC	Schwartz Criterion
UNDP	United Nations Development Program
WMS	Welfare Monitoring Survey

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Achieving sustainable economic growth with a focus on combating poverty has become the key development goal for governments around the world, as reflected in the Millennium Development Goals, in particular, Goal 1; *"eradicate extreme poverty and hunger"*. In relation to this objective, analysis of poverty aroused the interest of researchers, public authorities and international organizations. The specificities of developing economies, in particular the dualism between the urban and rural areas incites to identify the determinants of poverty with a view of designing policies and strategies to alleviate poverty that persists in most of these countries. Though literature on poverty measurements is relatively developed, yet studies about determinants or causes of poverty in Ethiopia are scarce due to lack of organized database.

Poverty in Ethiopia has many manifestations. The Human Development Index (HDI) for 2006, which takes life expectancy, adult literacy, primary schooling and per capita income as a basis ranked Ethiopia 170 out of the 177 countries. Ethiopia also ranks 92 out of 95 on the Human Poverty Index (HPI). Recent estimates suggest that about 31 million people (44%) live below the poverty line, which is equivalent to US45 cents or 3 Birr per person a day. Of every 1000 children born alive, 123 die before the age of five years (EDHS, 2005). There are 676 maternal deaths for every 100,000 live births, compared with 11 for every 100,000 in the UK (EDHS, 2011). The annual income for population of Ethiopia is $141 per person, compared with $37,000 in UK. According to the country health statistical report (2009), the average life expectancy at birth is 55.4 years in Ethiopia, compared with 79 years in the UK (www.dfid.gov.uk).

In 2002, the percentage of population with access to safe and clean water was only 22% and it increased to 54% in 2011 (EDHS, 2011). The majority of households, 82%, use non-improved toilet facilities (91% in rural areas and 54% in urban areas) which in turn affect the health of the community. Although the average annual growth rate of the country was by double digit, accesses to health services are inadequate for the majority of the population. Besides, only 38% of women are literate, an increase from 29% in 2005 (EDHS, 2011).

To fight against poverty, the government of Ethiopia has designed a 'Poverty Reduction Strategy Paper' (PRSP). According to MoFED (2002), the overall objective of this strategy is to reduce poverty through enhancing rapid economic growth while at the same time maintaining macroeconomic stability. Poverty headcount ratio is projected to decline by about 10% by the end of the PRSP period (2004/2005) from its 1999/00 level of 44%. To attain the Millennium Development Goals (MDGs) by 2015, Ethiopia adopted again the 'Plan for Accelerated and Sustained Development to End Poverty' (PASDEP), the second poverty reduction strategy, covering the period 2005/06 to 2009/10. In light of this, it is stated that the government is committed to work towards meeting the MDGs of 2015. And achieving such MDGs requires the Ethiopian economy to grow in real terms by 5.7% per annum until 2015 to reduce poverty by half from its current level. In addition to the above strategies, the Growth and Transformation Plan (GTP) has been developed and under process to maintain rapid and broad-based economic growth and eventually to end poverty (MoFED, 2010).

All governments make poverty reduction part of their policy agendas. However, the most important question is how these governments are going to achieve the goal. This could not be adequately addressed unless we have information on the characteristics of the poor and how these characteristics determine poverty. On this background, addressing the issue in terms of exploring the causes and examining ways of alleviating such problem is important. This can be done at macro level or concentrate on particular areas or groups of the society. However, for the fact that poverty is ultimately a problem at an individual or household level, micro-level studies on poverty are appropriate (Coulmbe & McKay, 1993). Identifying the household-level determinants of poverty is, therefore, vital for improving the living standards of the society and for the development of the country in general. Undertaking such a study is an appropriate way of producing data for evidence based decision making in the country.

1.2 Statement of the Problem

Poverty reduction has remained a crucial policy debate worldwide. All over the world workshops, conventions and conferences are held by governments and other development organizations to deal with poverty. Despite this, poverty continues to afflict the lives of many people especially in the developing world.

At country level, Ethiopia is on track in achieving the first MDG of halving poverty by 2015. Figures released by the Central Statistical Agency (CSA) and MoFED (2010-11) showed that the incidence of poverty has been falling since 1995/1996. The incidence of poverty declined from 45.5% in 1995/1996 through, 44.2% in 1999/00 to 38.7% in 2004/05. The above statistics do not mean that Ethiopia has won the battle against poverty. Poverty alleviation therefore still remains the most important challenge facing the country.

Based on this background the current study aimed at identifying the correlates (determinants) of poverty in Ethiopia at household level. In Ethiopia, the existing literature on the determinants of poverty is populous with the models mostly on the national or to some extent on the disaggregated models for urban/rural areas. The present paper extends the existing literature on poverty in Ethiopia by modeling and determining the various socioeconomic and demographic household level indicators and factors responsible for the poverty using the latest Household Income, Consumption and Expenditure Survey (HICES) data set which was obtained from CSA. The study further extended the analysis to three sub-models for the national level, urban and the rural area of Ethiopia.

1.3 Objective of the Study

The general objective of the study is to seek the best set of variables that predict the household poverty status using easily measurable household characteristic and socio-demographic indicators. More specifically the study attempted to:

- construct poverty indices at national level (headcount index, poverty gap index and the squared poverty gap index);
- analyze the relationship between household's poverty level with age, educational level, sex, marital status, employment status of household head, place of residence, family size and dependency ratio at national level;
- analyze the relationship between household's poverty level with age, educational level, sex, marital status, employment status of household head, family size and dependency ratio at urban and rural level separately.

1.4 Research Hypotheses

The welfare of household in Ethiopia is *positively* correlated with educational level dummies (primary school completed, secondary school completed and college and above), employment status of the household head (being self-employed or employed in the formal sector), male headship, age of household head, number of working members/productive age and residing in urban. We further hypothesized that, the welfare of household to be *negatively* correlated with dependency ratio and large family size.

1.5 Significance of the Study

There are reasons why this study had been conducted: (i) the study could help policy makers and other relevant organizations to understand the correlates/determinants of poverty at household level in Ethiopia; (ii) could increase the understanding of the linkage between household and poverty; (iii) could become a source of literature (information) for other researchers who may want to conduct research on poverty in the future.

1.6 Limitation of the Study

Poverty is multidimensional, interlocked and complex phenomenon. The present study was carried out at National, Urban and Rural level with general objective of assessing the determinants of poverty at household level. In order to attain this aim the study made use of secondary data (HICES) which was obtained from CSA. However, the HICES include information regarding the household consumption expenditure dimension only. In this regard, the survey does not take in to account other factors that determine the household's poverty level. Thus, it ignores the multi-dimensional nature of poverty.

CHAPTER TWO

LITERATURE REVIEW

2.1 Conceptual Framework of Poverty

Poverty is not a simple phenomenon that we can learn to define by adopting the correct approach. It is a series of contested definitions and complex arguments that overlap and at times contradict each other. It is differently seen as a big phenomenon or small phenomenon, as a growing issue or a declining issue, as an individual problem or a social problem, as a country problem or a regional problem and as urban problem or a rural problem (Chaudhry, 2003).

The World Bank (2000) defines poverty as "deprivation in well-being". Such a definition, however, raises the important question of what 'well-being' is and how it should be measured. A range of approaches that address this issue exist, with an important conceptual distinction being that between the 'welfarist' and 'non-welfarist' approaches (Ravallion, 1994). The 'welfarist' approach assesses well-being solely on utility information, derived from the preferences of individuals. The non-welfarist approach, on the other hand, bases the assessment of well-being (welfare) on the attainment of certain basic achievements, such as food, clothing and shelter.

The distinction between the two approaches can be illustrated by considering the case of two individuals, in which the first is much more deprived (among other things) in food, clothing, shelter and medical attention than the second, but is nevertheless happier. The welfarist approach, noting that the first individual is happier than the second, will consider the first to be better off. But according to the non-welfarist approach, it is the second individual who is enjoying a higher well-being.

Sen (1979) argues that the neglect of non-utility information makes welfarism too restrictive. The non-welfarist approach answers to this criticism by concentrating on the satisfaction of basic needs deemed necessary for a good standard of living. Thus, the identification of specific forms of commodity deprivation (both absolute and relative) becomes central, and the well-being of individual's is assessed by such measures as income, nutrition and health. Ravallion, however, argues that '... (A) nagging worry about these approaches has been arbitrariness in deciding what commodities matter and (when necessary) how one should value one against another...' (Ravallion, 1994).

Another conceptual definition of poverty and well-being, which rejects both utility and commodity based measures, is found in the seminal works of Sen (1992). Sen (1985) argues that well-being depends on what kind of life a person is living, and what he/she is succeeding in 'doing' or 'being'. Thus, well-being is seen from the perspectives of 'functionings' and 'capabilities'. 'Functioning' is an achievement and 'capability' is the ability to achieve. 'Functioning' is related to the state of existence of a person such as whether a person is well nourished, clothed, educated or participates in society without shame. Capability, on the other hand, has to do with an individual's freedom in the choice of their life and 'functionings'. It follows that a poor person may be considered as one with low capabilities.

Despite the different conceptions of poverty and well-being highlighted above, most empirical studies exclusively consider the satisfaction of material needs by defining a basket of goods necessary to sustain a minimum standard of living. As a result, income and consumption expenditure have been the preferred and most widely used measures of well-being. While such an approach has the benefit of empirical tractability, it ignores the multi-dimensional nature of poverty.

In countries where reliable income data can be found, income has often been used to conduct poverty and welfare analysis. However, the generally preferred indicator of welfare has been consumption expenditure, in part because of the volatility of income. Income may fluctuate in an unpredictable manner, making it a 'noisy' indicator of welfare. Consumption tends to be less volatile than income because consumption smoothing opportunities such as saving, borrowing, and community based risk sharing are available to the poor. This suggests that current consumption, rather than current income, is a better indicator of both current and long term standard of living (Deaton, 1997).

Another consideration in poverty analysis relates to the fact that measures of consumption are generally available at the household level. To have a clear picture of the degree of deprivation requires that appropriate corrections be made to total household consumption or income to take account of the differing needs of households. Needs may differ because of variation in the size and/or composition of households. A common approach is to divide aggregate household consumption by household size to obtain consumption per capita.

2.2 Poverty Lines

An accepted poverty line is the second condition next to the derivation of real household per capita consumption required to estimate poverty and welfare indicators. A poverty line is country specific and this level of income or expenditure varies from one country to another. Irrespective of countries, households or individuals with per capita consumption falling below this line are considered poor, however; and households with per capita consumption above this line are considered non-poor. Since poverty lines are cutoff points separating the poor from non-poor. There are two main ways of setting poverty lines-absolute and relative in the poverty analysis.

2.2.1 Absolute Poverty

Absolute poverty is claimed to be an objective, even a scientific definition, and it is based on the nation of subsistence. Subsistence is the minimum needed to sustain life, and so being below subsistence level is to be experiencing absolute poverty because one does not have enough to live on. In short, absolute poverty refers to the position of an individual or a household in relation to the minimum cost of food and a set of basic needs consistent with spending patterns of the poor. All those who are unable to satisfy these needs are considered as poor.

2.2.2 Relative Poverty

Relative poverty or inequality refers to the position of an individual or a household in relation to the average income and or expenditure. The view that poverty is a relative phenomenon has been further strengthened by Glabraith (1977) who argues, "People are poverty stricken when their income even if adequate for survival falls radically behind that of the community".

In this study, we are not concerned with the competition and superiority, of both forms of poverty. Since, poverty theorists advocated that absolute poverty concept is more relevant with the development problems of developing economies than the relative concept. So, it is more appropriate to estimate and analyze absolute poverty in a developing country like Ethiopia where the average level of resources is limited.

There are a number of different approaches to the determination of poverty line. The most common ones are *direct calorie intake, food energy intake and cost of basic need methods*. The direct calorie intake method defines poverty line as the minimum calorie requirement for survival. Therefore, individuals who consume below a predetermined minimum level of calorie intake are taken to be under poverty. This relates poverty to malnutrition. The limitation of this method is that the cost of acquiring such basic calorie requirement is not taken into consideration. Besides it overlooks the non-food requirements.

The other most popular method of setting poverty line that can overcome such problem is the Food Energy Intake (FEI) method. This method of setting a poverty line tries to find consumption expenditure or income level at which a person's typical food energy intake (nutrient intakes) is sufficient to meet a predetermined food energy requirement. Hence, in this method under nutrition is viewed as "food energy poverty". The method also aims to measure consumption or income poverty rather than under nutrition because it takes not only the nutrient intakes in relation to requirements but also the incomes or consumption expenditures.

There are different ways of estimating the expenditure needed to arrive at the stipulated requirements. A common practice is simply to calculate the average income or expenditure of a sub-sample of households whose estimated caloric intakes are approximately equal to the stipulated requirements. Another procedure is to use a regression of the empirical relation between food energy intake and consumption expenditure. By fixing the food energy intake cut-off, and then running a regression of calorie intake against consumption expenditures or income, one can find the consumption expenditure or income level at which an adult equivalent typically attain that food energy intake (Lipton and Ravallion,1993).

However, despite its simplicity in calculation, there are some difficulties in applying the method. Setting the food energy requirement is problematic because the requirements vary across individuals and over time for a given individual. Setting the requirement needs the assumption about the activity level that determines energy requirements beyond those needed to maintain the human body's metabolic rate at rest (Ravallion and Bidani, 1994). Hence, they argued that the FEI method is weak in terms of offering a consistent and robust poverty line.

The third method of setting poverty line is the Cost of Basic Needs (CBN) approach. The measurement of poverty line based on basic needs, dates back to the work by Rowntree (1901) when he attempted to construct poverty line based on cost for basic needs such as food, housing and clothing. According to this approach, poverty is a lack of command over basic consumption needs, and poverty line is the cost of those needs. In order to set this poverty line, one has to first define the food poverty line by selecting a basket of food items typically consumed by the poor in which case the quantity is determined in such a way that the given bundle meets the predetermined level of minimum caloric requirement. Then, this basket is valued at the relevant prices. To allow for the non-food expenditure, the food share of the poorest quartile divides the food poverty line. Therefore, it could be noted that this method gives a representative poverty line accounting for both food and non-food expenditures, besides it is consistent across regions unlike food energy intake method. In order to do this, adjustments for spatial and inter-temporal variations could be made to establish a poverty line that is consistent across regions, groups and periods.

2.3 Measurements of Poverty

Income or consumption is traditionally used to measures material deprivation. Especially consumption rather than income is viewed as the preferred welfare indicator because consumption better captures the long-run welfare level than current income. Consumption may better reflect households' ability to meet basic needs. Income is only one of the elements that allow consumption. Consumption reflects the ability of household's access to credit and saving at times when their income is very low. Hence, consumption reflects the actual standard of living (welfare). Consumption is better measured than income. In most developing countries, income report of households is likely to be understated compared to consumption expenditure report. Income is so erratic and seasonal that it may be very difficult for respondents to recall. Hence, many of the income poverty measures (such as the headcount ratio, poverty gap ratio, and the squared poverty gap ratio) use consumption rather than income in the conduct of poverty analysis.

Consumption to be an indicator of household's welfare, it has to be adjusted for difference in the calorie requirement of different household members. This adjustment could be made by dividing real household consumption expenditure by an adult equivalent scale that depends on the nutritional requirement of each family member. The adult equivalent scale must therefore be different for different age groups and the gender of adult members. Besides, household consumption may have to be adjusted for differences in prices across regions and for different point of time to take care of the difference in the cost of basic needs across space and over time.

Total poverty here refers to an aggregate measure of poverty that takes into account both the food and non-food requirements. Here it is worth noting how poverty lines are established. The most widely used method of estimating poverty line is the *cost of basic needs method* because the indicators will be more representative and the threshold will be consistent with real expenditure across time, space and groups. First, the food poverty line is defined by choosing a bundle of food typically consumed by the poor. The quantity of the bundle of food is determined in such a way that the bundle supplies the predetermined level of minimum caloric requirement (2200 kilocalorie). This bundle is valued at local prices or at national average prices if the objective is to get a consistent poverty line across regions and groups. Then a specific allowance for the non-food goods consistent with the spending of the poor is added to the food poverty line. To account for the non-food expenditure, the food poverty line is divided by the food share of the poorest quartile or quintile. There are several methods to measure poverty. This section briefly presents some of the measures.

1. Headcount Index or Poverty Rate (P_0)

The poverty rate is the share of the population whose consumption (or income) is below the poverty line. This measure quantifies the share of the population that cannot afford to buy a basket of goods. When real per capita consumption, y_i, is ranked as:

$y_1 \quad y_2 \quad \dots \quad y_q \quad z \quad y_{q+1} \quad y_{q+2} \quad \dots \quad y_N$ the poverty rate, P_0 is given by:

$$P_0 = \frac{1}{N}\sum_{i=1}^{N} I(y_i < z) = \frac{1}{N}\sum_{i=1}^{q} i = \frac{N_q}{N} \tag{2.1}$$

Where N = total population

$I(.)$ = an indicator function taking a value of 1 (poor) if the bracketed expression is true, and 0 (non-poor) otherwise.

y_i = welfare indicator (per capita consumption)

z = poverty line

N_q = number of poor in the population

2. **Poverty Gap Index (P_1)**

The second measure of poverty is the Poverty Gap Index, P_1. It is the average, overall people, of the proportionate gaps between people's living standards and the poverty line (as a proportion of the poverty line). It is also called the depth of poverty index. Mathematically, the Poverty Gap Index is defined as:

$$P_1 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{z-y_i}{z}\right) I(y_i < z) = \frac{1}{N}\sum_{i=1}^{q}\left(\frac{z-y_i}{z}\right) \tag{2.2}$$

Where the variables are defined as in equation (2.1)

The poverty gap index measures the degree to which the mean consumption of the poor differs from the established poverty line (depth of poverty).

3. **The Squared Poverty Gap Index (P_2)**

It is the average of the squared relative gaps. P_2 is defined similar to the Poverty Gap Index except that the poverty gaps are squared, thus giving the highest weighting to the largest poverty gaps. The squared poverty gap index captures differences in consumption levels among the poor. This measure is also called the severity of poverty index. Taking the previous notations, P_2 can be defined as:

$$P_2 = \frac{1}{N}\sum_{i=1}^{q}\left(\frac{z-y_i}{z}\right)^2 \quad (2.3)$$

4. **The Foster-Greer-Thorbecke (FGT) Poverty Index**

The Headcount Index, the Poverty Gap Index and the Squared Poverty Gap Index belong to a family of poverty measures known as the Foster-Greer-Thorbecke Index. These are referred to as decomposable poverty measures. A poverty measure is said to be decomposable if the poverty measure of a group is a weighted of the poverty measures of the individuals in a group. The general formula for the FGT class of poverty measures is:

$$P_\alpha = \frac{1}{N}\sum_{i=1}^{q}\left(\frac{z-y_i}{z}\right)^\alpha ;\alpha \geq 0 \quad (2.4)$$

The parameter α is a measure of the sensitivity of the index to poverty and the poverty line z. Larger values of α put higher weight on the poverty gaps of the poorest people. By setting $\alpha = 0$, equation (2.4) reduces to a Headcount Index (P_0). If $\alpha = 1$, equation (2.4) becomes a Poverty Gap Index, aggregating the proportionate poverty gap, which shows the shortfall of the poor's income from the poverty line, expressed as an average over the whole population.

In Ethiopia, the methods described above were first applied in the context of the 1995/96 Poverty Analysis Report. This was based on the cost of 2,200 kcal per day per adult food consumption with an allowance for essential non-food items. The food and total poverty lines used since 1995/96 in the country are 648 and 1,075 birr at national average prices, respectively (Table 2.1). To use these poverty lines and compute poverty indices, the per adult consumption expenditure has been updated by deflating all food and nonfood consumption items by spatial price indices (disaggregated at the regional level relative to national average prices) and temporal price indices (relative to 1995/96 constant prices).

To calculate the 1999/00 and 2004/05 poverty indices, first the nominal values of per adult food and non-food consumption items were deflated by the spatial price indices (disaggregated at regional level relative to national average prices) and temporal price indices (relative to 1995/96 constant prices) to arrive at real per adult consumption. Second the 1,075 Birr poverty line is applied to real per adult household consumption expenditure in order to calculate head count, poverty gap and squared poverty gap indices. To compute the 2010/11 poverty indices, the 1995/96 poverty line has to be computed at 2010/11 prices. To do so groups of consumption items defined in 1995/96 that generate 2200 kilocalories are valued at 2010/11 national average prices in order to obtain food poverty line of 2010/11. Then this food poverty line is divided by the food share of the poorest 25% of the population to arrive at the absolute poverty line for year 2010/11. The food and absolute poverty lines for 2010/11 are determined to be Birr 1985 and 3781, respectively.

Table 2.1: Total (Absolute) and Food Poverty Line in Birr

	1995/96	2010/11
Kilocalorie per adult per day (kcal)	2200	2200
Food poverty line per adult person per year (Birr)	648	1985
Total poverty line per adult person per year (Birr)	1075	3781

Source: An Interim Report on Poverty Analysis Study, 2010-11

5. The Human Poverty Index (HPI)

The HPI is a non-income measure of poverty. The HPI is a measure of poverty that is increasingly being used by the United Nations Development Program (UNDP) in its human development reports. It is related to the Human Development Index (HDI) in that it measures deprivations in the three basic dimensions of human development that is captured in the HDI. These dimensions are first, a long and healthy life and its corresponding deprivation used in the HPI is the vulnerability to death at a relatively early age, as measured by the probability at birth of not surviving to the age of 40. Second, knowledge and the deprivation from this HDI dimension is the exclusion from the world of reading and communications, as measured by the adult illiteracy rate. Third, a decent standard of living and the derivation used in the HPI is a lack of access to overall economic provisioning, as measured by the unweighted average of two indicators, the percentage of the population without sustainable access to an improved water source and the percentage of children under weight for age (UNDP, 2005). The HPI is calculated as follows:

$$HPI = \left[\frac{1}{3}\left(P_1^{\alpha} + P_2^{\alpha} + P_3^{\alpha}\right)\right]^{1/\alpha} \tag{2.5}$$

Where: P_1 = the probability at birth of not surviving to age 40;

P_2 = adult illiteracy rate;

P_3 = un-weighted average of population without sustainable access to an improved water source and children under weight for age.

2.4 Empirical Literature

The poverty literature shows, a large number of empirical studies conducted on identifying the correlates of poverty. Comparison of poverty among households with different characteristics, such as gender of the household head, literacy, schooling, family size and other related variables are reviewed in this section from different countries.

The effects of different economic and demographic variables on the probability of a household being in poverty in Costa Rica was analyzed by Rodriguez and Smith (1994) using a logistic regression model to estimate. The authors found that the probability of being in poverty is higher, the lower the level of education and the higher the child dependency ratio, as well as for families living in rural areas.

There is considerable evidence of a strong negative correlation between household size and consumption (or income) per person in developing countries. The poor devote a high share of their income to goods such as food, tap water, cooking utensils, firewood and housing etc. Lanjouw and Ravallion (1995) test the robustness of the relationship between poverty and household size using Pakistan Integrated Household Survey (PIHS) and results confirm the negative relationship between household size and poverty, as the size of household increases the probability of being poor will increase.

McCulloch and Baulch (1998) have investigated poverty dynamics in rural Pakistan using a unique five-year panel data set from the second half of the 1980s. Their results confirm that while the incidence of income poverty in the panel is high, with between one-fifth and one-third of households in any year having incomes below the poverty line, turnover amongst the poor is also rapid. Conventional poverty status (Logit) regressions show that the probability of a household being in poverty is increased by its household size, the dependency ratio and district of residence but decreased by secondary education, land, the value of livestock and other assets owned. The age and sex of the household head together with basic education did not, however, alter a household's poverty status. Household size was found to increase the probability of entering poverty and decrease the probability of exiting poverty. This effect is consistent with the effect of this variable in standard poverty status regressions. However, neither the dependency ratio not district dummies, which were important in the poverty status regressions, have much impact on the probability of entry and exit from poverty.

The determinants of poverty in Uganda by using logistic regression model was examined by Adebua *et al* (2002).This study shows that household with better educated heads are less likely to be poor and large households are more likely to be poor. This confirms that the larger the household size, the poorer the household is. This is because the large number of household members would likely be children who are unproductive and yet they take a big proportion of household income in terms of schooling requirement, medical attention, food and clothing.

The household level determinants of poverty in Kenya were examined using both binomial and polychotomous logit models of poverty analysis by Geda, et al (2001). The study shows that poverty status is highly correlated with the level of education, household size and type of occupational activity and it is most prevalent in rural areas. Specifically, poverty falls as the level of education increases; it rises with household size and with engagement in agricultural activities. These effects persist in both binomial and polychotomous models used to determine poverty. In particular, extreme poverty falls rapidly as education increases and as farm households' shift to non-agricultural activities. Size of landholding does not emerge as a major determinant of hard-core poverty in the sample analyzed.

Goaied and Ghazouani (2001) identified factors contributing to poverty in rural and urban areas in Tunisia. They used cluster of household data for their research. Probit and logit models were used to estimate their model, assuming that twenty percent of the Tunisian population was poor, according to the Census data for this research year period. Per capita consumption expenditures were used as the welfare level indicator of the household, and controlling for the heterogeneity of clusters in the household survey. The sample contained both urban and rural households. Most of the variables measured socio-demographic characteristics of the household, as well as variables representing the time frame and geographic location of the household. The main differences between rural and urban poverty rate could be attributed to education of household head, the child dependency ratio, the ratio of male to female employees in the household, the socio-professional category of the head, the geographic location of the family residence, and the share of the food budget spent for cereal products. A female headship of the household was a factor but only in urban and not in rural areas.

2.5 Empirical Works in Ethiopia

The literature dealing with poverty in Ethiopia is limited, reflecting the lack of an appropriate and reliable household survey data that would allow the comparison of welfare across time. Alemu et al (2011) employed a probit regression for poverty determinants using primary data on some selected rural part of Ethiopia. Findings of this study show that family size, land and livestock holdings, diversification in crop production, engagement in non-farm activities and utilization of microfinance services to be important correlates.

Bogale et al (2005) study aimed to investigate the determinants of rural poverty in Ethiopia. They used the logit model to identify determinants of poverty using data of a one-year rural household survey conducted in three rounds in three districts of Ethiopia during the 1999/2000 cropping season (Bogale, 2002). The results indicate that entitlement failures understood as lack of household resource endowments to crucial assets such as land, human capital and oxen.

Sepahvand (2009) used data from the 1997 Ethiopian Rural Household Survey (ERHS) to study the determinants of rural poverty using the Foster-Greer-Thorbecke model. He found that the incidence of rural poverty was high in villages that have lower conditions for agriculture. These finding imply that poverty reduction could be achieve through effective policies toward improving the conditions for agriculture in the rural areas. Moreover, examination of the connection between different socioeconomic characteristics and poverty indicated that households consisting of household heads with higher age and availability of farmland were relatively less poor. But, households where the households head at most completed primary school suffer from most incidence of poverty.

Abbi Mamo (1997) used multivariate analysis to analyze the determinants of standard of living Addis Ababa based on the first round Ethiopian Urban Household Survey (EUHS) which was conducted in 1994. A multinomial logit model was used to estimate the probability of being in poverty depending on household-specific economic and demographic explanatory variables. The study showed that education, access to credit, employment status, gender, marital status and food shortage were significant determinants of poverty.

Tassew Woldehanna (2008) estimated econometric models of household welfare (OLS, Probit and Tobit models) to see the effect of household characteristics and other variables on household welfare. At micro level, agricultural potential, natural resource endowment, education, household size and access to credit, food market and economic infrastructure were found to be the main correlates of poverty in Ethiopia.

The studies reviewed above analyzed the different determinants of poverty applying different methodologies. However, there is a need of identifying whether there are changes of poverty determinants and need to update the poverty profile in Ethiopia using the latest HICE survey data set. The current study aims to fill this gap.

CHAPTER THREE

DATA SOURCE AND RESEARCH METHODOLOGY

3.1 Data Source

The data used in this study to analyze poverty were taken from the 2010-11 Household Income, Consumption and Expenditure Survey (HICES) for Ethiopia. The survey was conducted from 8 July 2010 through 7 July 2011 which covered both rural and urban areas of the country. The core objective of the 2010-11 HICE survey is to provide statistical data that enable to understand the consumption expenditure dimension of poverty.

For the purpose of representative sample selection, the country was divided into three broad categories, i.e., rural, major urban centers and other urban areas. Therefore, each category of a specific region, in most cases, was considered to be a survey domain for which the major findings of the survey are reported. However, Harari and Dire Dhawa had rural and urban categories, only; while Addis Ababa had only urban areas divided into10 sub-cities considered as survey domain or reporting levels.

In the first two categories, namely the rural and major urban, a two stage stratified sampling technique was implemented whereby the Enumeration Areas (EAs) were considered as a Primary Sampling Unit (PSU) and the households were considered as the Secondary Sampling Unit (SSU). The EAs were selected using the Probability Proportional to Size (PPS), size being the number of households obtained from the 2007 Population and Housing Census while the households were systematically selected from the fresh list of households within the EA made during the survey.

On the other hand, for the other urban category, a three stage stratified sampling technique was utilized. In this case, the urban centers, EAs and households were used as a PSU, SSU and the Tertiary Sampling Unit (TSU) respectively. Here, the PSUs and SSUs were selected using the PPS while the selection of households follow the same approach as described earlier. At country level, a total of 864 EAs and 10,368 households (12 households per EA) were selected to represent rural and a total of 1,104 EAs and 17,664 sample households (16 households per EA) were selected for urban domains, specifically, 576 EAs and 9,216 household and 528 EAs and 8,448 households to represent major urban and other urban areas, respectively. However, with respect to households, only 197 households were not covered by the survey. At the end it was possible to obtain very clean data from 27,835 households.

3.2 Poverty Line Assessment in Ethiopia

In order to assess the welfare level, one might look at the household income data as a possible indicator of the household welfare level. The use of income data is not preferred because of the fact that income is often understated and provides biased estimates for the poverty analysis. The use of real total consumption per adult equivalence instead of income is favorable due to the fact that the expenditures actually represent the permanent income of a household. The minimum expenditures required to maintain a specific level of wellbeing is set as a threshold or called poverty line. The assessment of the minimum level of wellbeing is not arbitrary rather the cost of a basket of essential consumption goods is taken as a reference category. To control the poverty line for varying household sizes, the real total expenditure of each household is divided by the corresponding household size.

The second approach is the calorific approach that takes into account only food items for poverty line determination. The official poverty line of Ethiopia is calculated by selecting a basket of food items to meet the minimum required level of calorie intake of 2,200 calories per day per person and the cost of such a basket at the prevailing prices is calculated to set the minimum amount required for meeting the recommended nutritious level for a single person. This level is scaled up with some pre-specified multiple to obtain the final poverty threshold per capita.

3.3 Poverty Line Used in this Study

Instead of calculating a new poverty line to be used in this study we decided to use the official poverty line for Ethiopia which has been constructed by MoFED in 2010-11.

Table 3.1: Total (Absolute) Poverty Lines in Birr 1995/6-2010/11 per year

	1995/96	2010/2011
Poverty Line	1075	3781

Source: An Interim Report on Poverty Analysis Study, 2010-11

3.4 Definitions of Variables Used in the Study

The dependent variable and independent variables that were considered to affect the status of household's poverty were selected based on experiences from the available similar studies and the available data on the subject.

Table 3.2: List of Variables and their Description Used in the Study

Dependent Variable	
Pov_Stat	*Poverty status based on per capita consumption (0 = Not-poor, 1 = Poor)*
Independent (Explanatory) Variables	
SEX	Sex of the Household Head (0 = Male, 1 = Female)
AGE	Age of the Household Head (in year)
FSZ	Number of Household Members (Family Size)
FSZSQ	Family Size Squared
AREA	Place of Residence of Household (0 = Urban, 1 = Rural)
NWOR	Number of Working Members/Productive Age (between 15 and 64 years inclusive)
AGRL	Household Having Agricultural Land (0 = No, 1 = Yes)
DEPR	$\text{Dependency Ratio} = \dfrac{\text{People of (Age 14 and Below + Age 65 and Above)}}{\text{People Above Age of 15 and Below Age of 64}}$
EDLEV	Head of the Household Has No Education (NSCH = 0)** Head of the Household Completed Elementary School (CMPE = 1) Head of the Household Completed Secondary School (CMPS = 2) Head of the Household Completed College/University & Above (CCUA = 3)
MARST	Head of the Household is Single (SINGLE= 0)** Head of the Household is Married (MARRIED = 1) Head of the Household is Divorced/Widowed (DIVSEW = 2)
EMPST	Head of the Household is Employed in Informal Sector (INFOE = 0)** Head of the Household is Employed in Formal Sector (FORME = 1) Head of the Household Head is Self-Employed (SELFE = 2)

Note: ** shows reference category

3.5 Method of Data Analysis

The multiple logistic regression was adopted in order to observe the effects of the independent variables (**X**) on the dependent variable, poverty status *(pov_stat)*. In this study a household is considered to be poor if its total consumption per capita is below the official poverty line, that is birr 3,781 per year.

$$pov\_stat = \begin{cases} 1, poor \\ 0, not - poor \end{cases}$$

Let π_i denote the conditional probability that the i^{th} household is below the poverty line. Thus, the model for which the outcome variable is binary and having more than one explanatory variable can be written as:

$$y_i = \pi_i + \varepsilon_i;\ i = 1, 2, ..., n \qquad (3.1)$$

where:

$$\pi_i = \frac{\exp(z_i)}{1 + \exp(z_i)} \qquad (3.2)$$

with $z_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi} = \mathbf{X'\beta}$. Here is n×1 vector of response having $y_i = 0$ if the household is not-poor and $y_i = 1$ if the household is poor, **X** is an n× (p+1) design matrix of explanatory variables, **β** is a (p+1) ×1 vector of parameters, **ε** is also an n×1 vector of unobserved random errors. The quantity π_i is the probability for the i^{th} covariate satisfying the important requirement $0 \le \pi_i \le 1$. Then, the log-odds of having $y = 1$ for given is modeled as a linear function of the explanatory variables as:

$$E(y_i) = \pi_i = \ln\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi} \qquad (3.3)$$

The function $\pi_i = \dfrac{\exp(\mathbf{X'\beta})}{1+\exp(\mathbf{X'\beta})}$ is known as logistic function. The most commonly used method of estimating the parameters of a logistic regression model is the method of Maximum Likelihood (ML) instead of Ordinary Least Square (OLS) method. Mainly for this reason the ML method based on Newton-Raphson iteratively reweighted least square algorithm becomes more popular with the researchers (Ryan, 1997). The sample likelihood function is, in general defined as the joint probability function of the random variables whose realizations constitute the sample. Specifically, for a sample of size n whose observations are $(y_1, y_2, \ldots, y_n)$, the corresponding random variables are $(Y_1, Y_2, \ldots, Y_n)$.

Since the Y_i's are a Bernoulli random variable, the probability mass function of Y_i is

$$f_i(y_i) = \pi_i^{y_i}(1-\pi_i)^{1-y_i} \text{ where } y_i = 0, 1;\ i = 1, 2, \ldots, n \tag{3.4}$$

Since y_i's are assumed to be independent, the log-likelihood function L ($\boldsymbol{\beta}$) is defined as:

$$L(\boldsymbol{\beta}) = \sum_{i=1}^{n}\left[y_i \ln\left(\frac{\pi_i}{1-\pi_i}\right)\right] + \sum_{i=1}^{n}\ln(1-\pi_i) \tag{3.5}$$

For convenience in multiple logistic regression models, the likelihood equations can be written in matrix notation as:

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \mathbf{X}'(\mathbf{y}-\boldsymbol{\pi}) \tag{3.6}$$

Where $\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & . & . & x_{1p} \\ 1 & x_{21} & x_{22} & . & . & x_{2p} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & x_{n1} & x_{n2} & . & . & x_{np} \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ \beta_p \end{bmatrix} \ \& \ \boldsymbol{\pi} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ . \\ . \\ \pi_n \end{bmatrix}$

Computer-intensive numerical search procedures are required to find the maximum likelihood estimates $\hat{\boldsymbol{\beta}}$ and hence $\hat{\pi}_i$, because the multiple logistic regression model computes the probability of the selected response as a function of the values of the predictor variables. There are several widely used numerical search procedures, one of these employs iteratively reweighted least squares algorithm. In this study, we shall rely on standard statistical software programs (SAS) to obtain the maximum likelihood estimates of parameters.

3.6 Model Selection Strategy for the Data

Model selection is a fundamental task in data analysis, widely recognized as central to good inference. It is frequently employed as a way to identify the model that is best supported by the data (referred to as the 'best model') among the candidate set. Once we have the full model (all variables added in), then we have to select which independent variables should be kept in the model. To do so, multiple automated variable selection methods have been developed. Out of the available methods, this study uses two approaches.

3.6.1 Likelihood Ratio Test (LRT)

Likelihood ratio test is a test frequently used to determine whether data support a full model over a reduced model. The full model is accepted as 'best' when the likelihood ratio (reduced model negative log-likelihood: full model negative log-likelihood) is sufficiently large i.e the difference is unlikely to have occurred by chance (i.e. $p < 0.05$). Let we begin with the full logistic model with response function:

$$\pi = [1 + \exp(-\mathbf{X}'\boldsymbol{\beta}_F)]^{-1} \tag{3.7}$$

Where: $\boldsymbol{\beta}_F = (\beta_0, \beta_1, \beta_2, ..., \beta_p)$ and $\mathbf{X}'\boldsymbol{\beta}_F = \beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi}$

The reduced logistic model has the response function:

$$\pi = [1 + \exp(-\mathbf{X}'\boldsymbol{\beta}_R)]^{-1} \tag{3.8}$$

Where: $\boldsymbol{\beta}_R = (\beta_0, \beta_1, ..., \beta_q), \mathbf{X}'\boldsymbol{\beta}_R = \beta_0 + \beta_1 x_{1i} + ... + \beta_q x_{qi}$ and $q < p$.

Now, we find the maximum likelihood estimates for both models and evaluate their likelihood functions. Let it be known that p and q are the parameters for two models. The hypothesis that will be tested is:

$H_0 : \beta_{q+1} = \beta_{q+2} = ... = \beta_p = 0$

H_a : *Not* H_0.

As with all hypotheses testing, we need a test statistic for LRT, which is denoted as G^2.

$$G^2 = -2\ln\left[\frac{L(R)}{L(F)}\right] = -2[\ln L(R) - \ln L(F)] \tag{3.9}$$

$G^2 \sim \chi^2_{p-q}$ when n is large and under H_0.

The appropriate decision rule is if $G^2 > \chi^2_{\alpha}(p-q)$, reject H_0 and conclude that the full model is valid.

3.6.2 Information Criteria Approach

Another way to evaluate a model is to use information criteria metrics which attempt to quantify how well our model would have predicted the data.

3.6.2.1 *Akaike Information Criterion (AIC)*

The first information criteria to gain wide-spread acceptance as a model selection tool was the AIC. Akaike developed this AIC criterion, which takes model complexity into account, for the identification of an optimal parsimonious model from a class of competing models. It is given by:

$$AIC = -2LL + 2(p+1) \tag{3.10}$$

Where LL is the maximum log-likelihood and p is the number of parameters in the model. The second term is sometimes referred to as the penalty term, and adjusts to the size and complexity of the model. When the number of parameters increases, the first term becomes smaller. This is due to the fact that the more parameters there are, the more the chance of what is observed can happen. So, to adjust for this bias, AIC adds the term $2(p+1)$ to -2LL as a penalty for increasing the number of parameters (Hilbe, 2009). A small value of AIC means that the model is a better fit; however, this does not mean that the model is a good fit or a perfect model.

3.6.2.2 *Schwartz Criterion (SC)*

The Schwartz criterion also known as the Bayesian Information Criterion (BIC) is a model selection criterion designed to find the most probable model (from a Bayesian perspective) given the data. It is given by:

$$SC = -2LL + (p+1)\log(n) \tag{3.11}$$

Where LL is the maximum log-likelihood of the model, $(p+1)$ is the number of parameters in the model, and n is the number of observations in the data set. The second term is the penalty term and does the same as the AIC penalty term. When the number of parameters and data increase the -2LL decreases. The smaller the SC the better the model is for the selection process.

3.7 Model Adequacy Checking

After estimating the coefficients, there are several steps involved in assessing the appropriateness, adequacy and usefulness of the model.

3.7.1 Hosmer-Lemeshow Test

The test divides subjects into deciles based on estimated probabilities, and then computes a chi-square from observed and expected frequencies. Then a probability (π) value is computed from the chi-square distribution with 8 degrees of freedom to test the fit of the logistic model. The hypothesis to be tested is:

H_0: the model fits the data very well versus

H_a: the model does not fit the data very well at α level of significance.

The Hosmer-Lemeshow (H-L) test statistic, $\hat{C}$ for this hypothesis is given by:

$$\hat{C} = \sum_{k=1}^{10} \frac{(O_k - E_k)^2}{V_k} \tag{3.12}$$

Where O_k is the observed number of events in the k^{th} group, E_k is expected number of events in the k^{th} group, and V_k is a variance correction factor for the k^{th} group. If the p-value related to H-L goodness-of-fit test statistic is greater than 0.05, we will not reject the null hypothesis that there is no difference between observed and model-predicted values, implying that the model estimates fit the data at an acceptable level. That is, well fitting models show non-significance on the H-L goodness-of-fit test, indicating that model-prediction is not significantly different from the observed values. This does not mean that the model necessarily explains much of the variance in the dependent, only that however much or little it does explain is significant. As the sample size gets large, the H-L statistic can find smaller and smaller differences between observed and model-predicted values to be significant.

3.7.2 Receiver Operating Characteristics (ROC) Curve

Another way to see if the model is good fit is by using the ROC curve. The ROC curve is a plot of the true positive rate (sensitivity) against the false positive rate (1-specificity) for the different possible cut-offs of a diagnostic test. The area under the ROC curve, which ranges from zero to one, provides a measure of the model's ability to *discriminate* between those subjects who experience the outcome of interest versus those who do not. The results of goodness-of-fit can therefore be summed up in a single number: the area under the ROC curve (often abbreviated to AUC). As a general rule:

If AUC = 0.5, this suggests no discrimination (predictions were no better than random guessing);

If $0.7 \leq AUC < 0.8$, this is considered acceptable discrimination

If $0.8 \leq AUC < 0.9$, this is considered excellent discrimination

If $AUC \geq 0.9$, this considered outstanding discrimination (Hosmer and Lemeshow, 2000).

3.7.3 Residual Analysis and Residual Plots

Residual analysis for logistic regression is more difficult than the linear regression models because the responses take on only values 0 and 1. Let y_i denote the number of "successes" for n_i trials at setting i of the explanatory variables. Let $\hat{\pi}_i$ denote the estimated probability of success for the model fit. Then, the estimated binomial mean $n_i \hat{\pi}_i$ is the fitted number of successes. Thus the i^{th} ordinary residual will assume one of the two values as:

$$\hat{\varepsilon}_i = \begin{cases} 1 - n_i \hat{\pi}_i \; if \; y_i = 1 \\ -n_i \hat{\pi}_i \; if \; y_i = 0 \end{cases} \quad where \; i = 1,2,\ldots,n \tag{3.13}$$

The ordinary residuals will not be normally distributed and, indeed their distribution under the assumption that the fitted model is correct is unknown. Plots of ordinary residuals against fitted values will generally be uninformative. Hence, the ordinary residual can be made more comparable by dividing them by the estimated standard error of y_i, which is known as Pearson residual denoted by pr_i and defined as:

$$pr_i = \frac{\hat{\varepsilon}_i}{\sqrt{n_i\hat{\pi}_i(1-\hat{\pi}_i)}} = \frac{y_i - n_i\hat{\pi}_i}{\sqrt{n_i\hat{\pi}_i(1-\hat{\pi}_i)}} \tag{3.14}$$

The Pearson residuals are directly related to the Pearson chi-square goodness-of-fit statistic. The Pearson residuals do not have unit variance since no allowance has been made for the inherent variation in the fitted value $n_i\hat{\pi}_i$. A better procedure is to further standardize the ordinary residuals by their estimated standard deviation that is called studentized Pearson residuals. The standard deviation is approximated by

$$\sqrt{n_i\hat{\pi}_i(1-\hat{\pi}_i)(1-h_{ii})} \quad where \ \mathbf{H} = \hat{\mathbf{W}}^{1/2}\mathbf{X}\left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\mathbf{W}}^{1/2}$$

h_{ii} is the i^{th} diagonal element of the n×n estimated hat matrix **H**, whereby in logistic regression it is called hat diagonal or Pregibon leverage and measures the leverage of an observation. More clearly leverage is a measure of the importance of an observation to fit of the model. Here, $\hat{\mathbf{W}}$ is the n×n diagonal matrix with elements $n_i\hat{\pi}_i(1-\hat{\pi}_i)$, **X** is the n× (p+1) design matrix. Then the studentized Pearson residuals spr_i are defined as:

$$spr_i = \frac{y_i - n_i\hat{\pi}_i}{\sqrt{n_i\hat{\pi}_i(1-\hat{\pi}_i)(1-h_{ii})}} = \frac{pr_i}{\sqrt{1-h_{ii}}} \tag{3.15}$$

Studentized Pearson residuals are primarily helpful in identifying influential observations and those build in information about the influence of a case, whereas Pearson residuals do not. More influence cases with high leverage result in high studentized residuals. Studentized Pearson residuals approximately follow the standard normal distribution for large $(n \geq 30)$ sample and it can be used as an approximate chi-square distribution.

Deviance residual is another type of residual. It measures the disagreement between any component of the log likelihood of the fitted model and the corresponding component of the log likelihood that would result if each point were fitted exactly. Since, the logistic regression uses the maximum likelihood principle; the goal in logistic regression is to minimize the sum of the deviance residuals. Deviance residuals can also be useful for identifying potential outliers or misspecified cases in the model. The deviance residual for the i^{th} observation is the signed square root of the contribution of the i^{th} case to the sum for the model deviance, for the i^{th} observation, and is given by:

$$D_i = \pm\left\{2\left[y_i \ln\left(\frac{y_i}{n_i \hat{\pi}_i}\right) + (n_i - y_i)\ln\left(\frac{n_i - y_i}{n_i(1-\hat{\pi}_i)}\right)\right]\right\}^{1/2} \tag{3.16}$$

Where the sign, + or -, is the same as the sign of $(y_i - n_i \hat{\pi}_i)$.

Like Pearson residual the square of each deviance residual measures the contribution of each binary response to the goodness-of-fit statistic (Hosmer and Lemeshow, 2000).

The change in the value of estimated coefficients is analogous to the measure proposed by Cook (1977) for linear regression. It is obtained as the standadized difference between $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}_{(-i)}$, where these represent the maximum likelihood estimates based on full data set and excluding the i^{th} case respectively and standardizing via the estimated covariance matrix of $\hat{\boldsymbol{\beta}}$. Thus, one step linear approximation is given as:

$$\Delta\hat{\boldsymbol{\beta}}_i = (\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(-i)})'(\mathbf{X}\hat{\mathbf{W}}\mathbf{X})(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(-i)}) = \frac{pr_i^2 h_{ii}}{(1-h_{ii})^2} = \frac{spr_i^2 h_{ii}}{(1-h_{ii})} \tag{3.17}$$

Such a useful diagnostic statistic is one that examines the effect of deleting single subject on the value of the estimated coefficients $(\boldsymbol{\beta})$ and the overall summary measures of fit, like Pearson chi-square (χ^2) statistic and deviance (D) statistic (Agresti, 2007).

Chapter Four

Results and Discussion

An overview of the results obtained in the study are presented and discussed in this chapter. The data analyses involved both descriptive and inferential statistics.

4.1 Descriptive Statistics

Descriptive statistics allow researchers to present data acquired in a structured, accurate and summarized manner. The descriptive statistics utilized in the current study analyze the data using bar chart, percentages, means and standard deviations. Besides, chi square tests were used to explore the association between categorical independent variables and whether the household is poor or not. Moreover, the t-test was used to know the mean variations between the poor and not-poor in terms of continuous explanatory variables.

4.1.1 Household Size and Poverty Status

The mean family size of the country, according to the survey statistics, is found to be 4.74 persons per household. The average household size for the poor and non-poor households with respect to their poverty status is indicated in Table 4.1.

Table 4.1: Family Size by Poverty Status

Poverty Status	Mean	St. Dev	Total	
Poor	5.77	2.158	7715	$t-value=-71.09$
Non-poor	3.71	2.200	20120	$p-value=0.0001$

Source: Own computation from the survey

According to the findings computed from the survey and summarized in Table 4.1, the average family sizes of the sampled poor and non-poor households are 5.77 and 3.71 persons respectively. It shows that the mean household size of the poor category is greater than the non-poor category. The analysis of t-test showed that there is a significant difference between poor and non-poor in terms of household family size since the p-value is 0.0001 which is much less than the 5% level of significance.

4.1.2 Dependency Ratio by Poverty Status

The average dependency ratio (DEPR) for the country is computed to be about 81, which mean that every 100 persons at economically productive age group were responsible to take care of themselves as well as additional 81 persons (children and aged population) based on the survey. Table 4.2 shows the summary of dependency ratio with poverty status of the survey population.

Table 4.2: Dependency Ratio by Poverty Status

Poverty Status	Mean	St. Dev	Total	
Poor	1.270	0.994	7,715	$t-value=-51.44$
Not-Poor	0.634	0.761	20,120	$p-value=0.0001$

Source: Own computation from the survey

The survey has also indicated high variation in dependency ratio for poor households than the non-poor (0.994 vs 0.761). The mean dependency ratios for poor and non-poor households were estimated to be 127% and 63.4%, respectively. The analysis of t-test also showed that there is a significant statistical difference between poor and non-poor in terms of dependency ratio. Therefore, the analysis of the survey result is enough to prove the hypothesis that families with relatively higher number of dependents would increase household poverty.

4.1.3 Age of Household Head by Poverty Status

The average age of the poor household heads (44.92 year) is significantly greater than that of the not-poor household heads (40.46 year). This may indicate access to livelihood asset and wealth will be limited as household head become older. However, standard deviation for poor households is relatively higher showing relatively higher dispersion from the mean age.

Table 4.3: Age of Household Head by Poverty Status

Poverty Status	Mean	St. Dev	Total	
Poor	44.92	15.94	7,715	$t-value=-22.26$
Not-Poor	40.46	14.54	20,120	$p-value=0.0001$

Source: Own computation from the survey

The analysis of t-test also implies that there is a significant statistical difference *(p*-value = 0.0001 < 0.05*)* between poor and non-poor in terms of age of household heads.

4.1.4 Sex of Household Head by Poverty Status

As it can be seen from Table 4.4, out of 27,835 counts of the household heads, 8,757 (31.5%) households are female-headed and 19,078 (68.5%) are male-headed households. With respect to poverty status, 1,971 (7.1%) female-headed households are categorized under poor and 6,786 (24.4%) female-headed households are belonging to non-poor category.

Table 4.4: Sex of Household Head by Poverty Status

	Poverty Status			
Sex	Not-poor	Poor	Total	
Male	13334 (47.9%)	5744 (20.6%)	19078 (68.5%)	$\chi^2 = 173.051$
Female	6786 (24.4%)	1971 (7.1%)	8757 (31.5%)	$p-value = 0.000$
Total	20120 (72.3%)	7715 (27.7%)	27835 (100%)	

Source: Own calculation from the survey

The result from the chi square computation showed that there is a significant relationship between sex of household head and level of poverty since p-value = 0.000 which is less than 0.05.

4.1.5 Poverty and Urban-Rural Location

Table 4.5 illustrates the household's poverty status disaggregated by place of residence based on the official poverty line of Ethiopia constructed by MoFED in 2010/11. The households level of poverty of the sample comprises 8.1% (n = 2,249) of urban and 19.6% (n = 5,466) of rural household heads are below the threshold (poverty line), and thus categorized as poor. The tabular presentation also depicts that the proportion of rural poor household heads are higher (19.6% vs 17.4% for rural non-poor household heads).

Table 4.5: Place of Residence of Household Head Disaggregated by Poverty Status

	Poverty Status			
Area	Not-poor	Poor	Total	
Urban	15264 (54.8%)	2249 (8.1%)	17513 (62.9%)	$\chi^2 = 5215.804$
Rural	4856 (17.4%)	5466 (19.6%)	10322 (37.1%)	$p-value = 0.000$
Total	20120 (72.3%)	7715 (27.7%)	27835 (100%)	

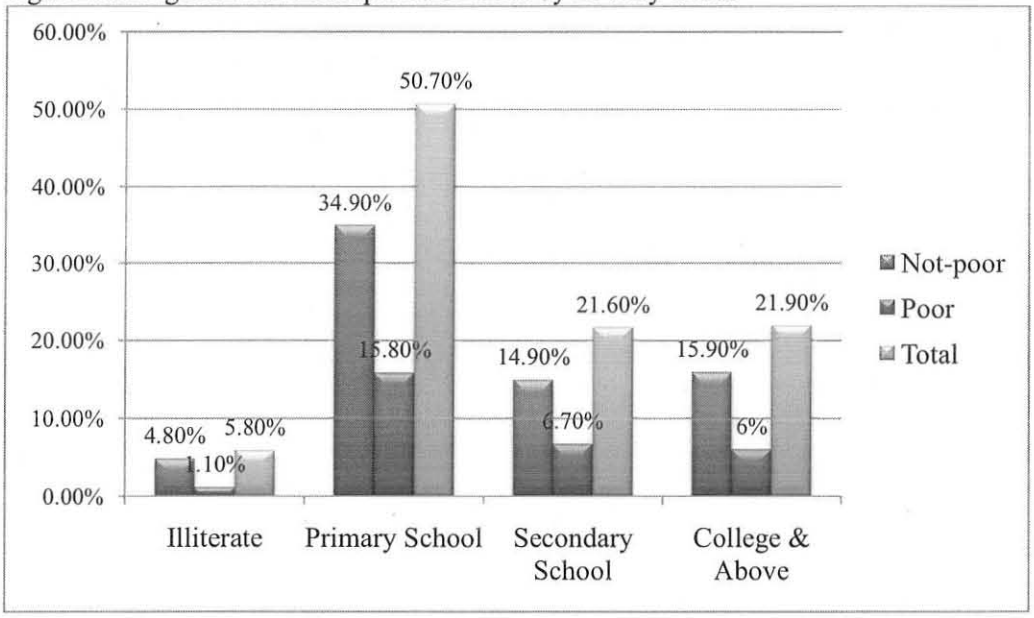
Source: Own calculation from the survey

Besides, the chi square value implies that there is a significant association ($p < 0.05$) between household head place of residence (AREA) and household's poverty status.

4.1.6 Educational Level of the Household Head by Poverty Status

As depicted in Figure 4.1, overall of the household heads reported, about 94.2% of the household heads in Ethiopia are literate with different level of schooling, the largest part of the sample population being in primary school (Grade 1 – Grade 8). In terms of literacy status, not-poor household heads do much better than poor household heads, accounting 65.7% and 28.4% respectively. In each level of schooling, most of the poor head households considered tend to be lower in number as compared to not-poor head households. The same as the not-poor household heads, most of the poor household heads tend to concentrate in primary school while the number of heads with school levels higher than secondary school is very small for poor heads.

Figure 4.1: Highest Grade Completed of Head by Poverty Status



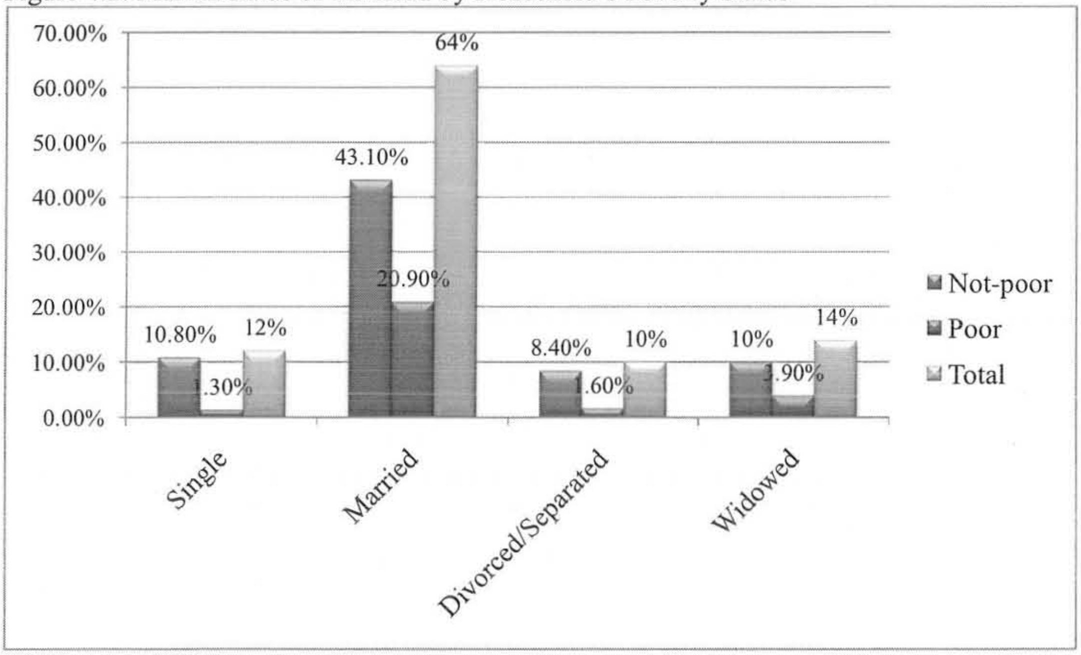
Source: Own computation from the survey (Pearson $\chi^2 = 79.857, p-value = 0.000$)

From the output, we can see that the Pearson chi square is 79.857 with p-value 0.000, indicating that there is sufficient evidence that educational level of household head and household's poverty status has a significant association at 0.05 level of significance.

4.1.7 Marital Status of Head by Poverty Status

Marital status is another socio-demographic variable to be considered. It can be seen from Figure 4.2, more than half of the reported household heads in Ethiopia, that is 64%, were married, followed by widowed (14%) and never married (12%) household heads.

Figure 4.2: Marital status of the Head by Household's Poverty Status

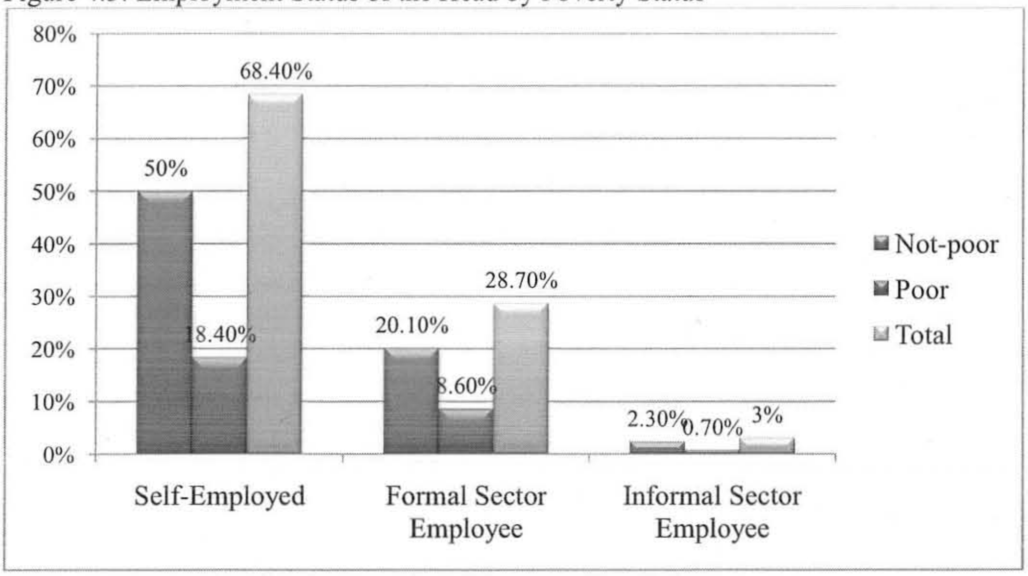


Source: Own computation from the survey (Pearson $\chi^2 = 909.288, p-value = 0.000$)

Among the poor household heads, it has been found that most of the poor heads, 20.90%, are married followed by widowed and divorced/separated poor heads accounting 3.90% and 1.60% of the poor household heads respectively. On the contrary, 10.80% of not-poor household heads, as opposed to 1.30% of poor heads, were single (never married). The chi square analysis resulted in a test statistic of 909.288 with p-value < 0.05. On the basis of this we rejected the null hypothesis and conclude that there is sufficient sample evidence to warrant rejection of the null hypothesis that marital status of household head is independent of the household's poverty level.

4.1.8 Employment Status of Household by Poverty Status

Figure 4.3: Employment Status of the Head by Poverty Status



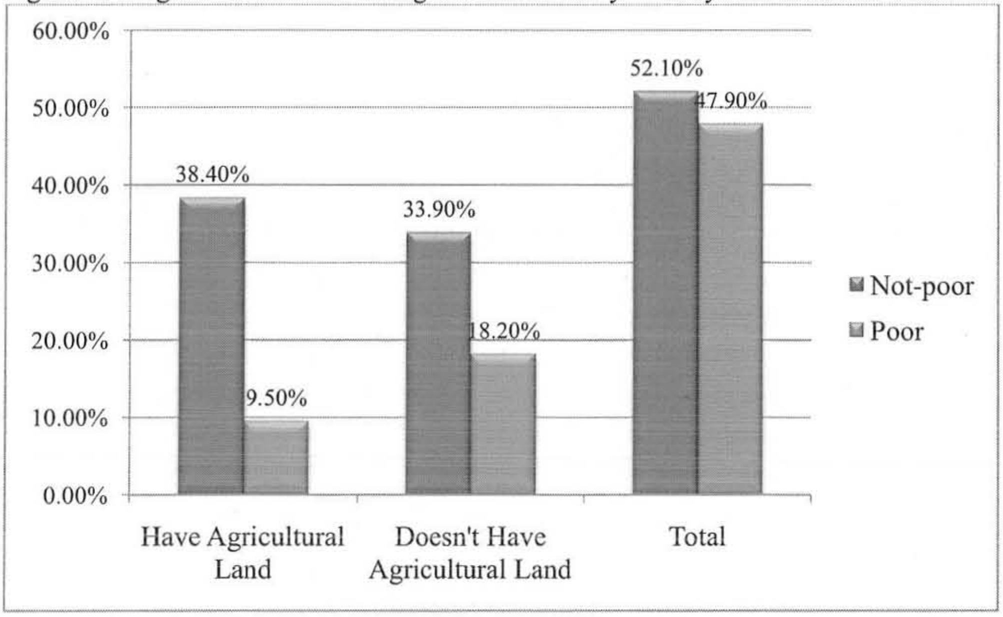
Source: Own computation from the survey (Pearson $\chi^2 = 29.187, p-value = 0.000$)

It can be viewed in Figure 4.3 that the majority of household head respondents 50% and 18.40% were self-employed for both not-poor and poor categories respectively. Besides, chi square test of independence was conducted to assess whether households head employment status was associated with the poverty status. The result showed that they were significantly dependent since p-value < 0.05.[1]

[1]Formal sector implies public sector, private sector and non-governmental organizations whereas the informal sector implies that house servant, unpaid family worker, apprentice (unpaid).

4.1.9 Agricultural Landholding of Household by Poverty Status

Figure 4.4: Agricultural Landholding of Household by Poverty Status



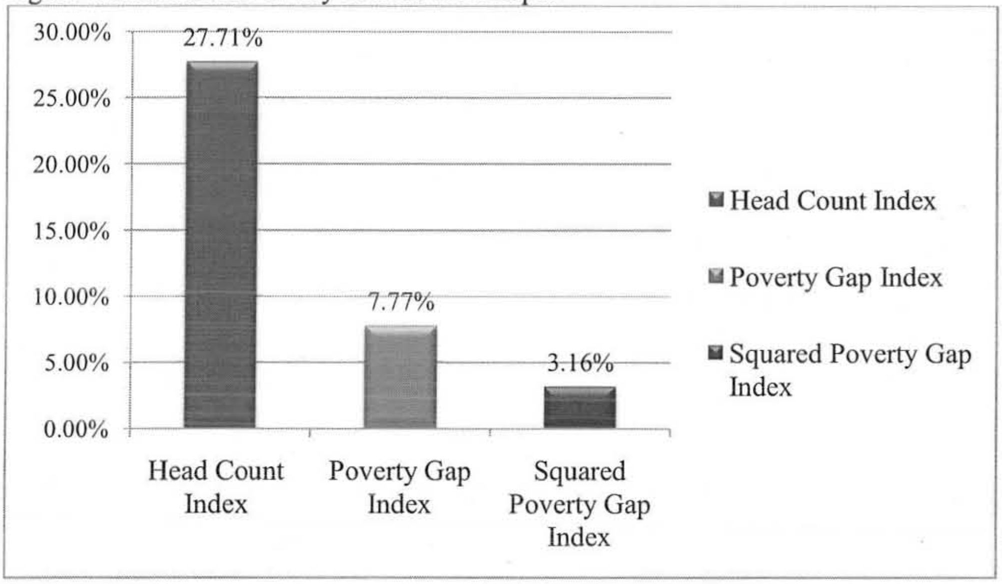
Source: Own computation (Pearson $\chi^2 = 779.152, p - value = 0.000$)

As depicted in Figure 4.4, the possession of agricultural (farming) land of the sample comprises 2655 (9.50%) and 10680 (38.40%) of poor and not-poor households respectively. However, 5060 (18.20%) and 9438 (33.90%) of poor and not-poor households does not have agricultural land. Further, we conclude that whether having agricultural land or not was statistically associated with the poverty status of the household since p-value = 0.000 is much less than the 0.05 level of significance.

4.2 Poverty Indices

The estimated poverty indices for Ethiopia using HICES (2010-11) are presented in Figure 4.5. Based on total poverty line, absolute headcount index stood at about 28% indicating that the percentage of the sampled population who was unable to meet the required minimum amount of calorie intake i.e. 2,200 kcal per person per day. In other words, this proportion of households could not attain the minimum amount of consumption (Birr 3,781) to satisfy the minimum calorie requirement per adult equivalent per year. Accordingly, the head count index ($\alpha = 0$) indicated that 28% of the sampled population in Ethiopia was below absolute poverty line.

Figure 4.5: Absolute Poverty Indices in Ethiopia



Source: Own computation from the survey

The poverty gap index ($\alpha = 1$) which captures the total proportional shortfall (i.e., the difference between per capita consumption and total poverty line and then divided by the total poverty line) is 7.77 %. This means total consumption needed to bring the entire population to above the poverty line or the minimum level of living is 7.77%. It indicates the percentage of consumption expenditure deficit the poor faces so as to uplift the poor from the poverty line. If one simply adds up the difference between the expenditure measure and poverty line for all those who were below, one would obtain the total money required to eliminate poverty. In similar fashion, the squared poverty gap index ($\alpha = 2$) in consumption expenditure, which aims to measure the severity of poverty, showed that 3.16 % of the sampled population was below the poverty line implying severe degree of inequality among the lowest quartile population.

4.3 Determinants of Poverty: Inferential Statistics

In order to explore the correlates of poverty with the variables thought to be important in explaining poverty a binary logistic regression model was estimated, with dependent variable being dichotomous variable of whether the household is poor (1) or not-poor (0) based on total expenditure of each household at National, Urban and Rural level. The explanatory variables considered in the analysis were: sex, age, educational level, marital status, employment status of household head, agricultural landholding, place of residence, number of working members, dependency ratio and size of household.

4.3.1 Model Selection for National Data

In the process of model selection, we shall try to adhere the principle of parsimony, the Likelihood Ratio Test (LRT) and information criterion. Let us begin by fitting the full model with all variables added in for each level (National, Urban and Rural area). The full printout of model at National level is shown in the following table.

Table 4.6: Logistic Estimate of Poverty at National Level (Full Model)

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Odds Ratio
INTERCEPT	-7.245	0.1527	2250.574	0.0001	.
DEPR	0.186	0.0368	25.706	0.0001	1.205
NWOR	-0.095	0.0384	10.878	0.0010	0.910
FSZ	1.017	0.0384	702.694	0.0001	2.766
FSZSQ	-0.050	0.0025	389.424	0.0001	0.951
AGE	-0.007	0.0011	48.636	0.0001	0.992
SEX	0.471	0.0401	137.893	0.0001	1.602
AGRL	-0.297	0.0503	34.925	0.0001	0.743
AREA	1.655	0.0359	2128.093	0.0001	5.234
DIVSEW	0.443	0.0789	31.521	0.0001	1.557
MARRIED	0.409	0.0448	83.274	0.0001	1.505
CMPE	-0.168	0.0400	17.724	0.0001	0.845
CMPS	-0.413	0.0650	40.342	0.0001	0.388
CCUA	-0.669	0.0860	60.528	0.0001	0.515
SELFE	-0.445	0.0495	81.040	0.0001	0.640
FORME	-0.208	0.0820	6.350	0.0120**	0.812

** Significant at 5% level of significance

After using the SAS software package, we were able to find the -2log-likelihood for the model, which was found to be 23431.280 with 15 degrees of freedom. Looking at the significance of individual variables (Pr > Chi-Square), all variables are significant at 5% level of significance in the model. The -2log-likelihood for the constant only model may be obtained by fitting the constant only model. Thus, refitting the logistic regression model excluding all independent variables the method yields -2log-likelihood of 32854.561 with 1 degree of freedom.

Now, we can use the Likelihood Ratio Test (LRT) to see if the coefficient of the variables removed are in fact equal to zero, i.e. $H_0 : \beta_1 = \beta_2 = \ldots\ldots\ldots = \beta_{14} = \beta_{15} = 0$, thus not adding additional information to the model (where $\beta_1, \beta_2, \ldots, \beta_{15}$ are the coefficients of DEPR, NWOR,...,FORME respectively). Thus, the value of the likelihood ratio test statistic is:

$G^2 = 32854.561 - 23431.280 = 9423.281$

For $\alpha = 0.05$, we have $\chi^2_{0.05}(15) = 24.996$. Since $G^2 = 9423.281 > 24.996$ and the p-value for the test is $P(\chi^2(15) > 9423.281) = 0.000$ which is significant at the $\alpha = 0.05$ level. Thus, we reject the null hypothesis which states that all coefficients are zero and conclude that at least one is different from zero.

Besides, the following table showcases for the null and a model with all significant predictor variables (full model). Each model had its AIC, SC, R-Square and number of parameters to be estimated.

Table 4.7: Model Selection Criteria

Model	**AIC**	**SC**	**R-Square**	**No. of Parameters**
Null	32855.561	32859.005	-	1
Full	23463.280	23502.392	0.3991	16

After examining each model, it would have seen that the full model had the lowest value of both criterion (AIC and SC) since lower values of these statistics indicate a better fitting model by adjusting for the number of explanatory variables and the number of observations used in the model. Furthermore, the R-square value for the full model is 39.91%. Therefore, the estimated logit model for household level determinants of poverty at national level is given as follows:

$$logit(\hat{\pi}_i) = -7.25 + 0.186depr - 0.095nwor + 1.02fsz - 0.05fszsq - 0.007age + 0.471sex$$
$$- 0.297agrl + 1.66area + 0.443divsew + 0.41married - 0.168cmpe - 0.413cmps - 0.669ccua$$
$$- 0.4376selfe - 0.21forme \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots Model - I$$

Another way to analyze the effects of independent variables upon the probability of being poor is by looking at the change of odds ratio as the independent variables change. Table 4.6 (the last column) shows the odds ratios for each independent variable at national level.

As can be seen in the table, the variables dependency ratio (DEPR), family size (FSZ), household head being female (SEX), marital status (DIVSEW and MARRIED) and living in the rural area (AREA) had odd ratios greater than one, which means that these variables are positively correlated with the probability of being poor. On the contrary, the variables number of working members (NWOR), agricultural landholding (AGRL), family size squared (FSZSQ), age of household head (AGE), having completed elementary education (CMPE), having completed secondary education (CMPS), having college education and above (CCUA), being self-employed (SELFE) and being household head employed in formal sector (FORME) all had odd ratios lower than one, which means that these variables are negatively correlated with the probability of being poor.

4.3.2 Model Selection for Urban Data

To develop appropriate model at urban level we might use the same procedure as that we develop for the national level. To know the 'best' household level correlates of poverty in urban area we begin by fitting the full model with all variables added in. Using the urban data, the output for a logistic regression model predicting the probability of being poor with the predictor variables is given in the following table.

Table 4.8: Logistic Estimate of Poverty at Urban Level (Full Model)

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Odds Ratio
INTERCEPT	6.189	0.1946	1011.465	0.0001	.
DEPR	0.288	0.0518	31.061	0.0001	1.335
NWOR	-0.202	0.0435	21.588	0.0001	0.817
FSZ	0.884	0.0518	291.178	0.0001	2.421
FSZSQ	-0.041	0.0033	161.371	0.0001	0.959
AGE	-0.031	0.0183	2.924	0.0870**	0.969
SEX	0.522	0.0542	93.016	0.0001	1.686
AGRL	-0.298	0.0504	35.187	0.0001	0.742
DIVSEW	0.056	0.0282	4.603	0.0042	1.058
MARRIED	-0.007	0.0702	0.011	0.9180**	0.993
CMPE	-0.391	0.0590	43.834	0.0001	0.676
CMPS	-0.398	0.1080	13.633	0.0002	0.671
CCUA	-0.829	0.1452	32.640	0.0001	0.436
SELFE	-0.669	0.0795	70.840	0.0029	0.512
FORME	-0.461	0.0982	22.038	0.0009	0.630

** Since 0.0870 and 0.9180 are greater than 0.05 thus, AGE and being MARRIED of household head are statistically insignificant.

The output for a model predicting household level correlates of poverty is given in Table 4.8 with -2log-likelihood of 11089.903. From the table, we can infer that the variables AGE and the dummy for marital status MARRIED were not significantly associated with likelihood of being poor or not (with the corresponding p-values 0.0870 and 0.9180 respectively).

To obtain the 'best' fitting model while minimizing the number of parameters, the next logical step is to fit a reduced model containing only those variables thought to be significant, and compare it to the full model containing all the variables. However, whenever a categorical independent variable is included (or excluded) from a model, all of its design variables should be included (or excluded); to do otherwise implies that we have recoded the variable (Hosmer and Lemeshow, 2000). For example, if we only include design variable DIVSEW for marital status, as defined in Table 3.2, then marital status is entered into the model as a dichotomous variable coded as divorced/widowed or not. Due to this reason the design variable MARRIED is kept in the model. If k is the number of levels of a categorical variable, then the contribution to the degrees-of-freedom for the likelihood ratio test for the exclusion of this variable will be k -1. Hence, refitting the logistic regression model excluding the non-significant variable (AGE only) we obtain the following.

Table 4.9: Logistic Estimate of Poverty at Urban Level (Reduced Model)

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Odds Ratio
INTERCEPT	-6.219	0.1920	1049.315	0.0001	.
DEPR	0.248	0.0289	73.751	0.0001	1.282
NWOR	-0.202	0.0435	21.588	0.0001	0.817
FSZ	0.875	0.0510	294.136	0.0001	2.399
FSZSQ	-0.041	0.0032	162.238	0.0001	0.960
SEX	0.526	0.0540	94.641	0.0001	1.692
AGRL	-0.113	0.0353	10.268	0.0014	0.893
DIVSEW	0.114	0.0182	39.028	0.0059	1.120
MARRIED	-0.007	0.0702	0.011	0.9180**	0.993
CMPE	-0.457	0.0576	63.166	0.0001	0.633
CMPS	-0.655	0.0878	55.728	0.0001	0.519
CCUA	-0.467	0.1065	19.301	0.0001	0.626
SELFE	-0.024	0.0072	11.021	0.0009	0.976
FORME	-0.020	0.0092	4.870	0.0273	0.979

**Since 0.9180 > 0.05 thus MARRIED is statistically insignificant.

The reduced model had a -2log-likelihood of 11092.691 with 13 degrees of freedom. Again, we can use the Likelihood Ratio Test (LRT) to see if the variable removed is not significant, i.e. $H_0 : \beta_5 = 0$, thus not adding additional information to the model (where β_5 is the coefficient of AGE). The results are as follows:

$$G^2 = 11092.691 - 11089.903 = 2.788$$

For $\alpha = 0.05$, we have $\chi^2_{0.05}(1) = 3.841$. Since $G^2 = 2.788 < 3.841$, we fail to reject H_0 and conclude that variable age of household head (AGE) is not statistically significant for improvement of the model. The *p*-value of this test was 0.0949.

Consider also the suggested model in urban area being the reduced by looking at the model selection criteria as shown below in Table 4.10.

Table 4.10: Model Selection Criteria

Model	**AIC**	**SC**	**R-Square**	**No. of Parameters**
Null	13428.268	13436.038	-	1
Full	11123.691	11232.476	0.2364	15
Reduced	11090.903	11222.999	0.2388	14

The AIC, SC and – 2log-likelihood (the lower value the better) indicate that the model with the selected covariates (reduced) is superior to the full and model with intercept only (Table 4.10). The fitted logit model for household level determinants of poverty in urban area is given as follows:

$$\begin{aligned} &logit(\hat{\pi}_i) = -6.22 + 0.248depr - 0.202nwor + 0.875fsz - 0.0412fszsq + 0.53sex - 0.113agrl \\ &+ 0.114divsew - 0.0072married - 0.46cmpe - 0.656cmps - 0.468ccua - 0.024selfe \\ &- 0.0203forme \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots Model - II \end{aligned}$$

4.3.3 Model Selection for Rural Data

To specify or identify a model for rural area, the same as the previous two cases let us start with the model which includes all predictor variables. The full SAS output of model for rural area data is provided in the following Table 4.11.

Table 4.11: Logistic Estimate of Poverty at Rural Level (Full Model)

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Odds Ratio
INTERCEPT	-3.483	0.1935	324.235	0.0001	.
DEPR	0.465	0.0515	81.521	0.0001	1.592
NWOR	-0.047	0.0381	1.559	0.2117**	0.953
FSZ	1.202	0.0582	426.614	0.0001	3.330
FSZSQ	-0.061	0.0039	243.167	0.0001	0.941
AGE	0.001	0.0014	0.439	0.5074**	1.001
SEX	0.414	0.0610	46.106	0.0001	1.513
AGRL	-0.312	0.0512	37.133	0.0006	0.732
DIVSEW	0.661	0.1044	40.181	0.0001	1.932
MARRIED	0.126	0.0559	5.056	0.0021	1.339
CMPE	-0.278	0.0540	26.535	0.0001	0.757
CMPS	-0.370	0.0859	18.550	0.0001	0.690
CCUA	-0.602	0.0764	62.149	0.0001	0.547
SELFE	-0.779	0.0685	129.335	0.0001	0.459
FORME	0.054	0.0781	0.486	0.4865**	1.056

** NWOR, AGE, and FORME are statistically insignificant at 5% level of significance.

As we can see from the table above, for each parameter estimate, a standard error was estimated along with a Wald Chi-Square statistic, p-value for the Wald Chi-Square statistic and the corresponding odds ratio. The full model had a -2log-likelihood of 12016.248 with 14 degrees of freedom. From Table, it can be seen that the variables NWOR, AGE and the dummy for employment status FORME were not significantly associated with likelihood of being poor or not (with p-values 0.2117, 0.5074, and 0.4865 respectively). At this point, we would have to run a new logistic regression with statistically non-significant explanatory variables eliminated (NWOR and AGE), to refine the estimate of the outcome variable poverty status *(pov_stat)*.

Even if the variable FORME is statistically not significant (p-value = 0.4865 > 0.05) it stay in the model since FORME is a design variable for employment status of household head as indicated in Table 3.2.

Table 4.12: Logistic Estimate of Poverty at Rural Level (Reduced Model)

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Odds Ratio
INTERCEPT	-3.443	0.1781	373.900	0.0001	.
DEPR	0.285	0.0518	30.402	0.0001	1.331
FSZ	1.048	0.0476	484.145	0.0001	2.852
FSZSQ	-0.057	0.0038	224.813	0.0001	0.945
SEX	0.368	0.0599	37.759	0.0001	1.445
AGRL	-0.459	0.0506	82.210	0.0001	0.632
DIVSEW	0.647	0.1045	38.316	0.0001	1.909
MARRIED	0.702	0.0618	129.305	0.0001	2.011
CMPE	-0.286	0.0534	28.769	0.0001	0.751
CMPS	-0.610	0.0754	65.478	0.0001	0.543
CCUA	-0.398	0.0820	23.575	0.0001	0.676
SELFE	-0.800	0.0526	231.121	0.0001	0.449
FORME	-0.130	0.0556	70.570	0.0001	0.877

Within the framework of a statistical model, a set of data supports one statistical hypothesis better than the other if the likelihood of the first hypothesis, on the data, exceeds the likelihood of the second hypothesis. From the above SAS output (Table 4.11) which fits the full model, we see that -2log-likelihood = 12016.248. From the above SAS output (Table 4.12) which fits the reduced model, we see that -2log-likelihood = 12020.988. Therefore likelihood ratio test statistic is (G^2) = 12020.988 - 12016.248 = 4.74.

For $\alpha = 0.05$, we have $\chi^2_{0.05}(2) = 5.991$. Since $G^2 = 4.74 < 5.991$. In this case, the full model does not fit the data significantly better than the reduced model, and we infer that the two variables (NWOR and AGE) should be dropped from the model. The p-value of this test was 0.0934.

In addition, the following table showcases for the null, full and a model with all significant predictor variables (reduced model). Each model had its AIC, SC, R-Square and number of parameters to be estimated.

Table 4.13: Model Selection Criteria

Model	AIC	SC	R-Square	No. of Parameters
Null	14272.243	14279.485	-	1
Full	12046.665	12076.870	0.2622	15
Reduced	12041.321	12073.166	0.2599	13

AIC and SC serve as measures for comparing different models and hence the smaller the information criterion, the better model. Based on the model selection criteria as in Table 4.13, above, it is proposed that the best model among the three is the reduced model. The fitted logit model for household level determinants of poverty in Rural area is given as follows:

$$logit(\hat{\pi}_i) = -3.44 + 0.286depr + 1.048fsz - 0.057fszsq + 0.368sex - 0.459agrl + 0.647divsew$$
$$+ 0.703married - 0.2867cmpe - 0.61cmps - 0.398ccua - 0.8selfe - 0.131forme \ldots\ldots\ldots\ldots Model - III$$

4.4 Model Diagnostics

Once the model has been identified and estimates of the parameters involved have been found, we proceed to check if our model is adequate. In model diagnostics we are concerned with analyzing the quality of the model that we specified and estimated (Model-I, Model-II and Model-III).

4.4.1 Goodness-of-Fit of the Models (Hosmer-Lemeshow Test)

The goodness-of-fit measures how effectively the model describes the response variable. Now, we can test the reduced model to see if it is a good fit. The H-L test is implemented in SAS and the summary of computer output is given in the following table for models identified above (Model-I (National), Model-II (Urban) and Model-III (Rural)).

Table 4.14: Hosmer - Lemeshow Goodness-of-Fit Test

	National	Urban	Rural
H-L test statistic ($\hat{C}$)	2.223	2.6469	4.0756
P-value	0.9734	0.9545	0.8509
No. of observations	27827	17507	10320

For $\alpha = 0.05$, we have $\chi^2_{0.05}(8) = 15.507$.

Since $\hat{C}_{National} = 2.223, \hat{C}_{Urban} = 2.6469$ and $\hat{C}_{Rural} = 4.0756$ are all less than the tabulated value 15.507, we do not reject the null hypothesis, and conclude that the fit of the model is considered to be adequate for the three models. Thus, the goodness-of-fit test with *p*-values 0.9734, 0.9545 and 0.8509 indicate that there is insufficient evidence to claim that the models do not fit the data adequately. If the *p*-value is less than our accepted α-level (5% in this case), the test would reject the null hypothesis of an adequate fit. So our models fit the data well.

4.4.2 Classification Table and ROC Curves

In order to assess the predictive power of the models, a classification table of correct and incorrect predictions was constructed, based on the predicted probability of being poor for each data. A probability greater than or equal to 0.5 was interpreted as a prediction of a household being poor, while a probability less than 0.5 is interpreted a prediction of a household being not poor. Table 4.15 shows the classification table for the models. In this table, "D" represents the number of poor households in the sample while "~D" represents the number of not-poor cases in the sample. The symbol "+" represents the number of households predicted as poor by the model while "-" represents the number of not-poor cases predicted by the model.

Table 4.15: Classification Table of Correct and Incorrect Predictions for National [Urban] Rural

Classified	True		
	D	~D	Total
+	4272 [613] 4252	1974 [1001] 1859	6246 [1614] 6111
-	3443 [1636] 1214	18138 [14257] 2995	21581 [15893] 4209
Total	7715 [2249] 5466	20112 [15258] 4854	27827 [17507] 10320

	National	**Urban**	**Rural**
Sensitivity	55.30%	27.26%	77.79%
Specificity	90.10%	93.44%	61.70%
Positive predictive value	68.40%	37.98%	69.58%
Negative predictive value	84.05%	89.71%	71.16%
False + rate for true ~D	9.82%	6.56%	38.30%
False – rate for true D	44.63%	72.74%	22.21%
False + rate for classified +	31.60%	62.02%	30.42%
False – rate for classified -	15.95%	10.29%	28.84%
Correctly classified	80.53%	84.94%	70.22%

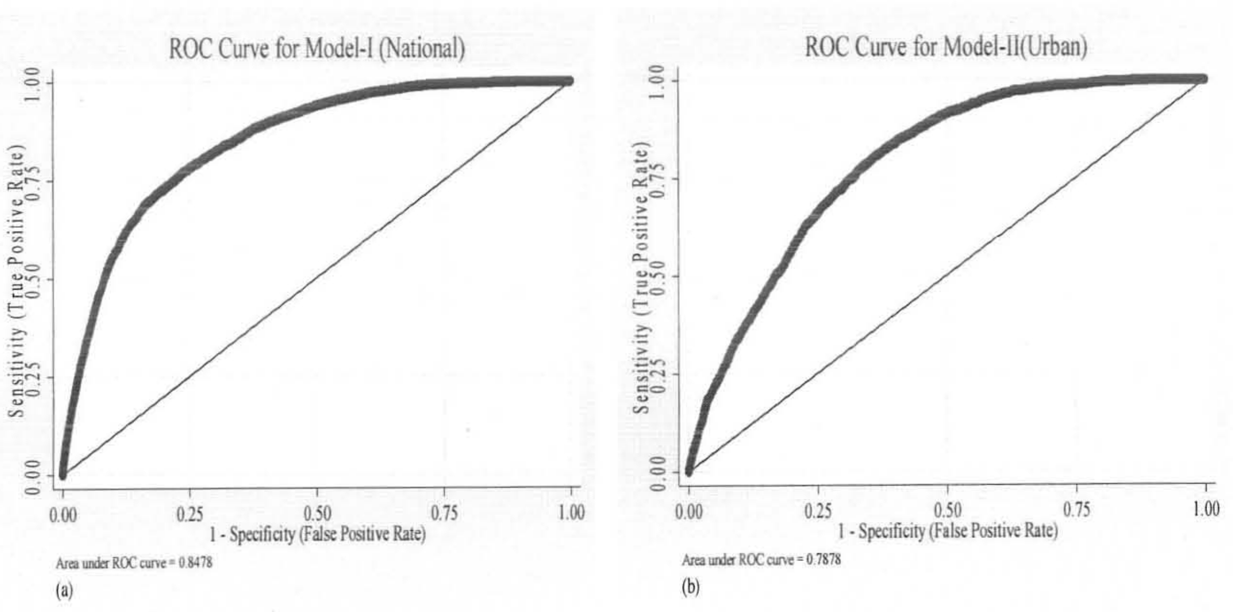
As can be seen in the Table, the models sensitivity rate (percent of poor cases correctly predicted by model) are 55.3%, 27.26% and 77.79%, while the models specificity rate (percent of non-poor cases correctly predicted by the model) are 90%, 93.44% and 61.70% for national, urban and rural respectively.

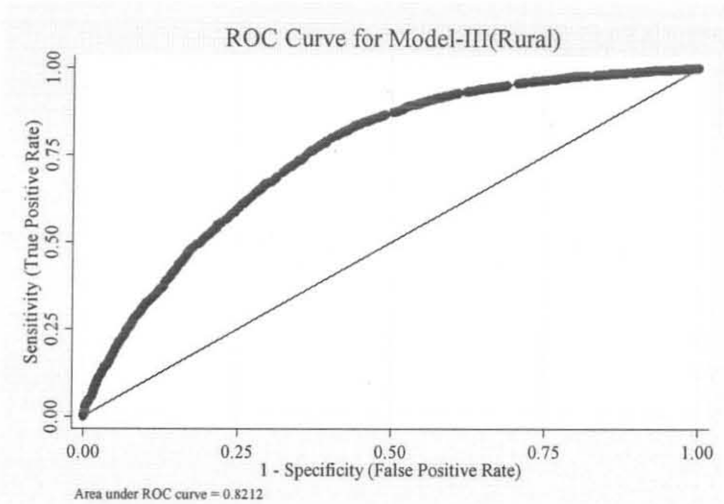
The false positive rate for households classified as poor by the model at national level is 31.6 percent, which means that 31.6% of the number of households predicted as poor by the model are in fact non-poor. The false negative rate for households classified as not poor by the model is 15.95%, which means that 15.95% of households predicted as not-poor by the model are in fact poor. Similarly, 62% and 30.42% of households are classified as poor by urban and rural model respectively since they are in fact not-poor.

The positive predictive value rate of the national model is 68.4 percent, which means that 68.4 percent of the total number of predicted poor households is in fact poor. Negative predictive rate is 84 percent, meaning that 84 percent of the total number of not-poor cases predicted by the model is in fact not-poor. As a whole, the National model correctly predicts 80.53% of cases where as 84.94% and 70.22% of the households were predicted correctly by urban and rural model respectively.

Sensitivity and specificity rely on a single cut point to classify a test result as positive. A more complete description of classification accuracy is given by the area under the ROC curve. If we look at the ROC curves, in Figure 4.6, which has an $AUC_{National} = 0.8478$, $AUC_{Urban} = 0.7878$ and $AUC_{Rural} = 0.8212$. This means that the models provide a good fit to the corresponding data.

Figure 4.6: ROC Curves for Model-I, Model-II and Model-III



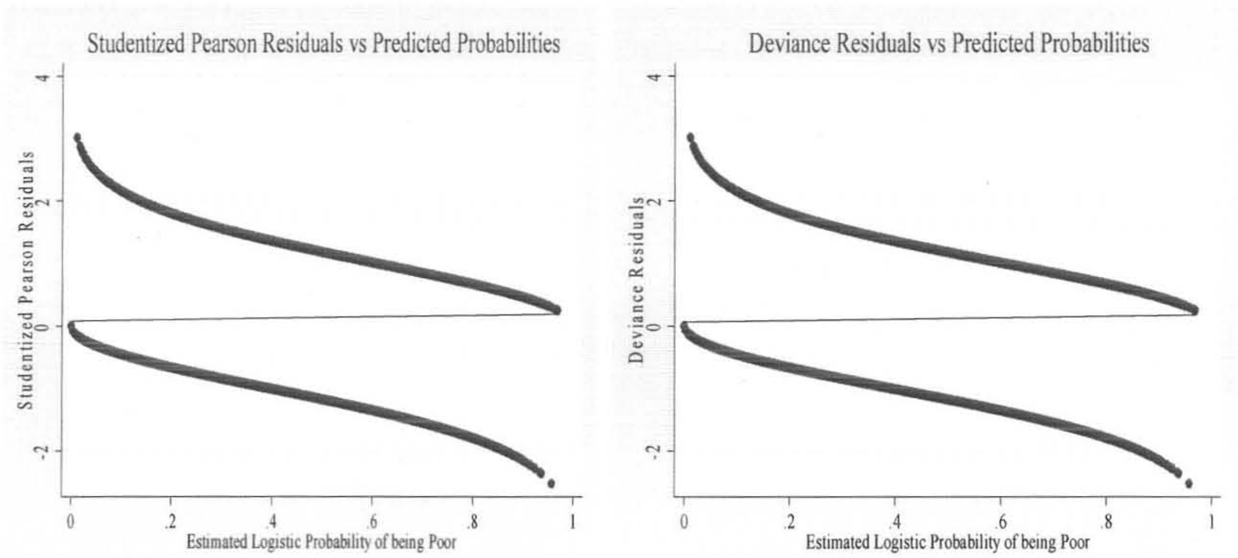


(c)

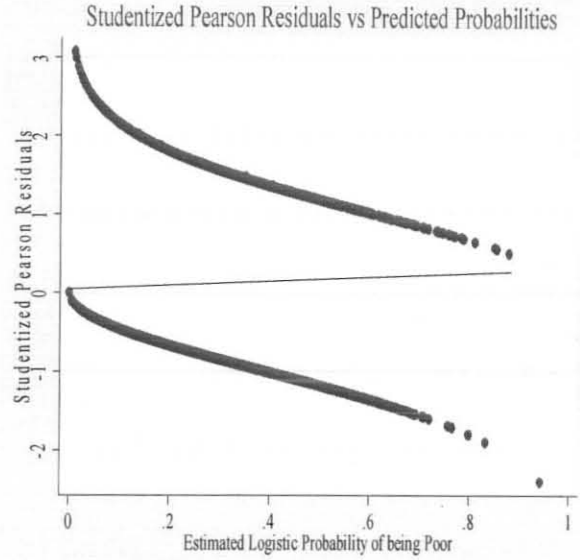
4.4.3 Diagnostic Plots

We can assess how great an influence a particular individual point has on the coefficients of the logistic regression model, and any lack of overall fit in all model. Any points with unduly high influence should be investigated to ensure that the data has been recorded correctly. The effect of removing the influential point on the model should also be investigated, and if necessary the results presented both with and without the influential point(s). One way of looking at the model adequacy is to graph studentized Pearson and deviance residuals against predicted probabilities.

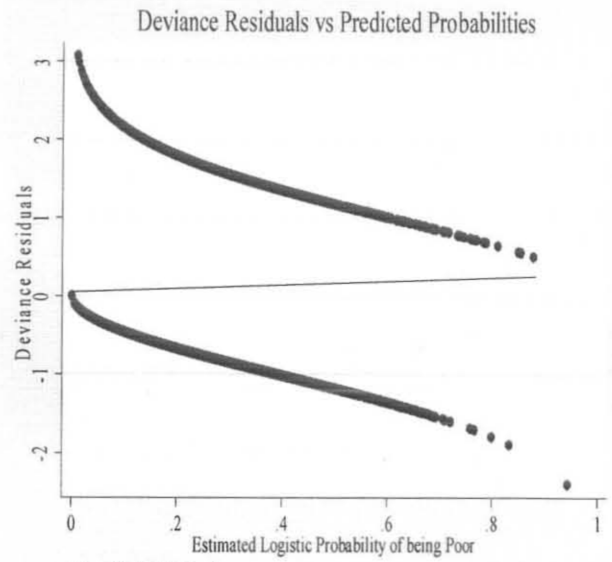
Figure 4.7: Standardized Residuals plotted against Estimated Logistic Probability of being Poor



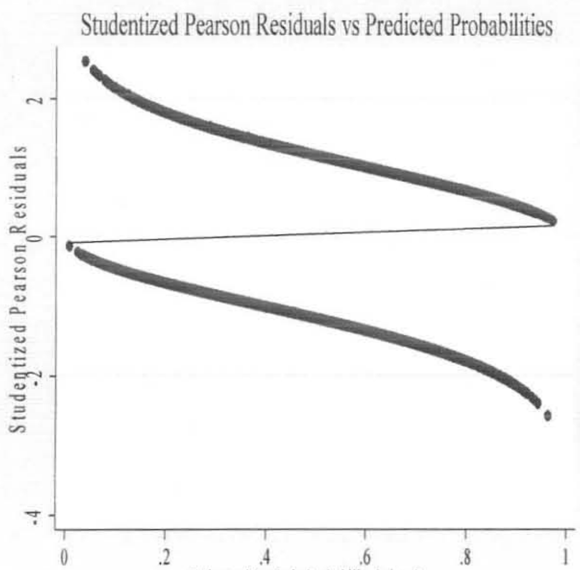
(a) For Model-I (National)
(b) For Model-I(National)



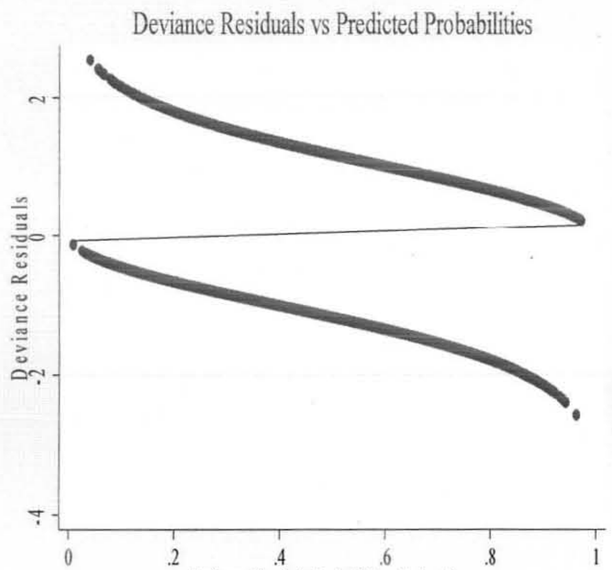
(c) For Model-II(Urban)



(d) For Model-II(Urban)



(e) For Model-III(Rural)



(f) For Model-III(Rural)

The lowess smooth of the studentized Pearson residuals and deviance residuals are demonstrated in Fig 4.7. In Fig 4.7 (a) - (f), the studentized Pearson residuals and deviance residuals are plotted against the estimated logistic probability respectively. In all cases, the lowess smooth approximates a line having zero slope and intercept. Any significant departure from this suggests that the model may be inadequate and potential outliers may have dramatic impact on the fit of the model (Sarkar *et al*, 2011).

The diagnostic test results for detection of outliers and influential values are presented in Table 4.16 using DFBETA (the change in the logistic regression parameter estimates when the observation is deleted) for National, Urban and Rural data.

Table 4.16: Delta Beta Statistics for the Data at National, Urban and Rural Level respectively

	National		Urban		Rural	
	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum
Cook's influence statistics	.00000	.19128	.00000	.27649	.00002	.11987
Normalized residual	-2.01245	2.60324	-2.0837	2.52632	-2.1509	1.88707
DFBETA for constant	-.00499	.02532	-.00919	.04133	-.00713	.03330
DFBETA for DEPR	-.00498	.00444	-.01136	.00604	-.00339	.00127
DFBETA for NWOR	-.00243	.00358	-.00528	.00461	**	**
DFBETA for AGRL	-.00091	.00087	-.00162	.00139	-.00182	.00178
DFBETA for FSZ	-.01362	.00277	-.02102	.00535	-.01368	.00324
DFBETA for FSZSQ	-.00015	.00105	-.00027	.00160	-.00034	.00125
DFBETA for AGE	-.00004	.00006	**	**	**	**
DFBETA for SEX	-.00265	.00173	-.00334	.00330	-.00329	.00208
DFBETA for AREA	-.00120	.00091	**	**	**	**
DFBETA for DIVSEW	-.00451	.00489	-.01040	.01242	-.00737	.00739
DFBETA for MARRIED	-.00145	.00194	-.00341	.01231	-.01938	.02167
DFBETA for CMPE	-.00104	.00107	-.00208	.00204	-.00198	.00174
DFBETA for CMPS	-.00284	.00269	-.00706	.00600	-.00507	.00369
DFBETA for CCUA	-.00326	.00343	-.00565	.00934	-.00541	.00532
DFBETA for SELFE	-.00164	.00171	-.00408	.00297	-.00206	.00274
DFBETA for FORME	-.00277	.00230	-.00711	.00549	-.00226	.00381

The above table shows that the normalized residuals are within the interval of -3 and 3 implying that no outliers were detected at 5% level of significance in the three cases. Since an observation is called influential if it has notable effect on parameter estimates, Cook (1977) proposed that the influence diagnostic must be larger than 1 for an individual case to have an effect on the estimated coefficients. Influence diagnostic (DFBETAs) corresponding to the outlying cases is presented in Table 4.16.

However, the values themselves are not especially large with respect to 1 and suggest that no observation had overall impact on the estimated vector of regression coefficients. Hence, it can be concluded that no significant model inadequacy and presence of influential outliers are observed. Thus, the existing outliers detected by the residual plots are not so influential and may not have dramatic impact on the fit of the three models.

Now, the reduced models (Model-I, Model-II and Model-III) have passed through the examinations (model checking procedures), it is a time to discuss the emperical results in comparison with the avialable literature in Ethiopia.

4.5 Discussion of the Results

The data has been analyzed at three levels (national, urban and rural). The results are provided in the form of the regression coefficients, standard errors, Wald Chi-Square, *p*-values and the corresponding odds ratio on the probability of remaining poor. Various socioeconomic and demographic variables have been considered for modeling the determinants of poverty.

One of the most salient facts about poverty in developing countries was that it was higher in rural area than urban area. For example, in Ethiopia by constructing a poverty profile Tassew Woldehanna (2008) found that, poverty was higher (consumption was lower) in rural areas than in urban areas. Although there may be problems associated to determine the direction of causality, several variables might explain why poverty is higher in rural areas than in urban areas. Firstly, lack of infrastructure. Rural poverty is often a product of poor infrastructure that hinders development and mobility. Poor infrastructure hinders communication, resulting in social isolation among the rural poor, many of whom have limited access to media and news outlets. Such isolation hinders integration with urban society and established markets, which could result in greater development and economic security. Secondly, rural areas are heavily dependent on agricultural production, which in developing countries is characterized by low labor productivity and therefore low incomes.

Our own estimates using the logistic regression for the 2010-11 survey indicated that rural-urban location is statistically significant as a cause of poverty in Ethiopia. As shown in Table 4.6, the odd of being poor for a household located in a rural area are 5.234 times the odd of an urban household.

Ethiopia, like other developing countries is experiences a high population growth rate. This high growth accompanied by the high unemployment rate and low female labor force participation rate poses a serious threat to the households. High dependency ratio (DEPR) and larger family size (FSZ) contributed positively to the probability of becoming a poor household for the three models (Model-I, Model-II and Model-III). The coefficients for both of these variables are positive and significant at 5% level of significance. The coefficient of family size squared (FSZSQ) is however negative and significant, controlling for the fact that very large families can also have potential earners and can reduce the poverty through larger participation in the work force. However, this situation is not highly desirable due to the fact that the odds ratio are at a very low level of less than 1% in reducing the probability of being poor for Model-I, Model-II and Model-III (0.951, 0.960 and 0.945, respectively). The odd ratio of the variable dependency ratio (DEPR) shows a contribution of 20.5%, 28.2% and 33.1% in increasing the likelihood of being poor where as family size (FSZ) contributes 176.6%, 139.9% and 185.2% for Model-I, Model-II and Model-III respectively.

Number of working household members/productive age (NWOR) had a potential in reducing the probability of remaining in the poor household category for Model-I and Model-II. Although the magnitude of the effects of NWOR for the two models is not similar (-0.095 for Model-I and -0.2022 for Model-II), yet the direction of the effect is negative for these two models. However, in Rural area (Model-III) NWOR was not found to be significant in lowering the possibility of being poor. This may lead to the conclusion that there were a larger number of dependents and lesser per capita consumption in rural area of Ethiopia.

Sex of household head i.e being female (SEX) positively affected the likelihood of remaining poor for the three models. Several studies had discussed the phenomenon of the feminization of poverty, which is said to exist if poverty was more prevalent among female-headed households than among male-headed households. This situation might be due to the presence of discrimination against women in the labor market, or it might be due to the fact that women tend to have lower education than men and they are paid lower salaries. Using a probit model, Meron Asefa (2003) found that female-headed households are poorer and more vulnerable to poverty than male-headed households in urban Ethiopia.

Looking at the results of logistic regression estimated above (Model-I), we reached the same conclusion as Meron (2003) since the sign for sex of the head (SEX) is positive and statistically significant at 5% level of significance. Moreover, the odd of being poor for female-headed households were 1.602 (OR = 1.602 with p-value = 0.0001) times for male-headed households. Conversely, we can say that the odd ratio of being poor for those male-headed were 0.6242 (OR = 0.6242, given by the reciprocal of 1.602) times for those headed by female.

Similarly, sex of household head (SEX) caused an increase in the likelihood of being poor for both urban and rural area. The odd of remaining poor for female-headed households are 1.692 and 1.445 times for male-headed households in urban and rural area of Ethiopia respectively. The finding is coincide with the Model-I but the magnitude of the effect is lower for the rural area.

It is argued that poverty increases at old age as the productivity of the individual decreases and the individual has few savings to compensate for this loss of productivity and income. This is more likely to be the case in developing countries like Ethiopia, where savings are low because of low income. However, the relationship between age and poverty might not be linear, as we would expect that incomes would be low at relatively young age, increase at middle age and then decrease again. Therefore, according to life-cycle theories we would expect to find that poverty is relatively high at young ages, decreases during middle age and then increases again at old age.

In the case of Ethiopia based on 1999/00 Household Income, Consumption and Expenditure Survey (HICES) and Welfare Monitoring Survey (WMS), Tassew Woldehanna (2008) found that age of the household head (AGE) was relevant in explaining poverty. Using the 2010/11 HICES and the methodology developed above we reached at the same conclusion as Tassew and age of the head is statistically significant in explaining poverty at national level (Model-I). However, even if the coefficient of AGE is significant, as can be seen in Table 4.6 above, an increase of one year in the age head decreased the odd of being poor by only 0.8 percent (OR = 0.992). In addition to this, the variable AGE is not statistically significant in lowering the probability of being poor for Model-II and Model-III; i.e, in urban and rural areas of the country separately.

In this world, people are hired into an occupational hierarchy and progress within it according to their skills and abilities. Thus, it is possible that education can have a favorable effect on well-being by allowing income to increase, which in turn, would lead a better life. It is often the case that if head of household is highly educated, the descendents will also be likely to get higher education. Education was grouped into four categories ranging from illiterates to those who have attended higher education (college and above). The odds of being poor with education level elementary school (CMPE), secondary school (CMPS) and college and above (CCUA) was found to be 0.845, 0.388 and 0.515 times that of illiterate (no schooling-reference category) respectively, implying that household head with higher educational attainment (CCUA) exhibited a lower chance to be poor as compared to households with illiterate household head for Model-I. The effect of educational level on household poverty had a similar trend in urban and rural area as in the case of national level (see Table 4.9 and Table 4.12 respectively).

The result of marital status (for Model-I) indicate that divorced/widowed and married clients are 55.7% and 50.5% more likely to be poor respectively than single (never married-reference category) clients, implying that for those whose marriage ended because of death of a partner or due to some disagreement were found to have a significantly high likelihood of being poor in Ethiopia. The category of marital status (MARRIED) has an ambiguous sign for both models (Model-II and Model-III). In Model-II (urban area), it is negative and insignificant while in Model-III (rural), it is positive and significant.

Employment status of the household head was one of the determinants of household's poverty status. Self-employed household head (SELFE) were about 36% less likely to be poorer than those employed in informal sector (INFOE) which is the reference category. Household which are headed by a person who was employed in formal sector (FORME) were about 19% less likely to be poorer than those worked in informal sector (INFOE) at country level (Model-I). Similarly, the employment dummies are all significant for urban and rural area and show that the employment sector was one correlate that affected the probability of being poor household.

Last but not certainly the least, the ownership of agricultural land of household (AGRL) significantly contributed to lowering the possibility of being poor at national, urban and rural level. The results showed that households having agricultural (farming) land were 25.7%, 9.7% and 36.8% less likely of remaining in the poverty trap respectively. The possible reason might be that the most of the population majorly employed in agricultural sector; the agricultural sector therefore is a big sector of employment in rural area especially as compared to urban area of the country.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This study analyzed the micro level determinants of household poverty in Ethiopia, both Urban and Rural, using 2010-11 HICES. The analysis on poverty dynamics highlighted that the poverty predictors may revolve around the demographic characteristics of the household and the household head. This might result due to in countries like Ethiopia where traditionally the head of a household head was the only earning person in the household. On the basis of information obtained, which is analyzed and interpreted in chapter four the following conclusions are made.

- About 28 percent of households (headcount index) are below the absolute poverty line, which was 38.7 percent in 2004-05. This implies that Ethiopia is on the right track in achieving the MDG target of reducing poverty by half in 2015.
- The results of household survey data indicated that poverty was most prevalent and pronounced in rural area than the urban area of the country.
- Poverty in Ethiopia was relatively widespread in larger a household (family size) which was the cause for slow improvement in living standard of the people along with their worse livelihood conditions in both urban and rural area of the country. In other words having fewer children (< 15 year) and elders (> 64 year) in the household contributed to the reduction of poverty.
- Positive relationship was observed between the female headship of household and the poverty status (being poor) at national, urban and rural level of the country. Results showed that a household headed by female have more probability to fall in poverty than male headed.
- Raising the level of education of a household head had a clear effect in reducing the probability of poverty in the two areas. The probability of poverty drops by slight percentages as the level of education rose from one level to the next.
- The evidence also suggested that employment status of head of the household was important for reducing the incidence of poverty in Ethiopia. It was found that the head of household being employed in the formal sector or being self-employed will bring ways of exit from the poor household category at national, urban and rural level of the country.

- Agricultural landholding not only for rural but also for urban households, was negatively associated with the poverty status, that is, being poor.

5.2 Recommendations

The results underline the fact that poverty is multifaceted and several prolonged approaches are needed. On the basis of the analysis made in this study the following important recommendations have been suggested for poverty alleviation.

- As human capital development is an important precondition for poverty alleviation through its impact on production and development process, meaningful linkage between formal and non-formal education should be interwoven to cover the large number of people in rural areas of the country.
- Rapid growth of the population relative to resource generation and lack of employment opportunities are the cause of poverty in the country because high population growth adversely affects the per capita consumption. Therefore, government should make efforts to control the increasing population by creating awareness among the masses and encourage them for having smaller families to somehow mitigate the likelihood of being poor. Besides, special emphasis could be given on creation of job opportunities for lifting a society out of poverty.
- Government should take initiative to create women friendly environment and encourage them as a diligent workforce.
- Intensive use of agricultural land is an essential strategy for alleviation of poverty in the country.
- Finally, it is recommended for future researchers to look at the severity or depth of poverty, by considering ordinal type of dependent variable taking multiple poverty lines and possible explanatory variables found relevant in the literature.

REFERENCES

Abbi Mamo (1997). Modeling Poverty and Its Determinants in Addis Ababa: A Focus on Multinomial Logit Selection. *Ethiopian Journal of Economics* **5**(2): 1-36.

Adebua, A, F. N. Okurut and J. A. Odowee (2002). Determinants of Regional Poverty in Uganda, Research Paper 122. *African Economic Research Consortium, Nairobi.*

Agresti, A. (2007). *An Introduction to Categorical Data Analysis.* Wiley Series in Probability and Statistics, 2nd Edition, Florida.

Alemu, D., Hagedorn, K. and Korf, B. (2011). Extent and Determinants of Household Poverty in Rural Ethiopia: A Study of Six Villages. *Eastern Africa Social Science Research Review,* **27**(2): 21-46 (http://muse.jhu.edu/journals/eas/summary/v027/27.2.alemu.html).

Bogale, A. (2002). Land Degradation, Impoverishment and Livelihood Strategies of Rural Households in Ethiopia: Farmers Perception and Policy Implications. Institutional Change in Agriculture and Natural Resources. Shaker, Aachen.

Bogale, A., Hagedorn, K. and Korf, B. (2005). Determinants of Poverty in Rural Ethiopia. *Quarterly Journal of International Agriculture* **44** (2): 101-120.

Central Statistical Agency [Ethiopia] and ICF International. 2012. *Ethiopia Demographic and Health Survey 2011.* Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Agency and ICF International.

Central Statistical Agency [Ethiopia] and ORC Macro. 2006. *Ethiopia Demographic and Health Survey 2005.* Addis Ababa, Ethiopia, and Calverton, Maryland, USA: Central Statistical Agency and ORC Macro.

Central Statistical Authority [Ethiopia] and ORC Macro. 2001. *Ethiopia Demographic and Health Survey 2000.* Addis Ababa, Ethiopia, and Calverton, Maryland: Central Statistical Authority and ORC Macro.

Chaudhry, I. (2003). An Empirical Analysis of the Determinants of Poverty in Pakistan: A Case Study of Bahawalpur District with Special Reference to Cholistan, PhD Thesis, Pakistan.

Cody, R. and Smith, J. (1997). *Applied Statistics and the SAS Programming Language,* 4th Edition, New Jersey 07458.

Cook, R.D., (1977). Detection of Influential Observations in Linear Regression. *Technometrics,* **19**: 15-18.

Coulombe, H. and McKay, A. (1993). The Causes of Poverty, In Poverty, Inequality and Rural Development. Lioyd, T., and. Morrissey, W., Macmillan London.

Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Baltimore and Maryland: The John Hopkins University Press.

Foster, J., J. Greer and E. Thorbecke (1984). A Class of Decomposable Poverty Measure, *Econometrica*, **52**: 761-6.

Galbraith, J. K. (1977).*The Affluent Society*. London, Andre Deutsche.

Geda, A, N.D.Jong, G.Mwabu and M.S. Kimenyi (2001). Determinants of Poverty in Kenya: A Household Level Analysis, Discussion Paper, Institute of Social Studies Hang and Kenya Institute for Public Policy Research and Analysis.

Goaed, S. and M. Ghazouani (2001). The Determinants of Urban and Rural Poverty in Tunisia, *Discussion paper, Laboratoire d'Econométrie Appliquée (LEA), Faculté des Sciences Economiques et de Gestion de Tunis, Tunisia.*

Hilbe, J. M. (2009). *Logistic Regression Models.* Boca Raton: Chapman & Hall/CRC Press.

Hosmer, D. and Lemeshow, S. (2000). *Applied Logistic Regression,* 2nd Edition, John Wiley & Sons Inc., New York.

Lanjouw, P and M. Ravallion (1995). Poverty and Household Size. *The Economic Journal*, **105**(433), 1415-1434.

Lipton, M. and Ravallion, M., (1993). Poverty and Policy, Working Paper Series 1130, The World Bank, Washington D.C.

McCulloch, N. and B. Baulch (1998). Distinguishing the Chronically from the Transitorily Poor: Evidence from Rural Pakistan, Working Paper No. 97, Institute of Development Studies, University of Sussex, Brighton.

Meron Assefa (2003). Female-Headed Households and Poverty in Urban Ethiopia, MA Thesis, Addis Ababa.

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

Ministry of Finance and Economic Development (MoFED) (2002). *Ethiopia: Sustainable Development and Poverty Reduction Program*, Addis Ababa.

Ravallion, M. (1994). *Poverty Comparisons. Fundamentals of Pure and Applied Economics,* Volume **56**, Harwood Academic Press, Chur, Switzerland.

Ravallion, M., and Bidani, B. (1994). How Robust is a Poverty Profile?, *World Bank Economic Review*, **8** (1).

Rodriguez, A. and Smith, S. (1994). A Comparison of Determinants of Urban, Rural and Farm Poverty in Costa Rica. World-Development; **22**(3): 381-97.

Rowntree, B.S. (1901). Poverty: A Study of Town Life, Macmillan.

Ryan, T. (1997). *Modern Regression Methods.* Har/Dis Edn, Wiley, New York, USA, ISBN-10: 0471529125.

Sarkar, S., Midi, H. and Rana, S. (2011). Detection of Outliers and Influential Observations in Binary Logistic Regression. *Journal of Applied Sciences* **11** (1): 26-35.

Sen, A.K. (1979). Issues in the Measurement of Poverty. *Scandinavian Journal of Economics,* **81**: 285-307.

Sen, A.K. (1985). Commodities and Capabilities, Amsterdam: North Holland.

Sen, A.K. (1987). The Standard of Living. Cambridge, Cambridge University Press.

Sen, A.K. (1992). Inequality Reexamined, Cambridge, Mass: Harvard University Press.

Sepahvand, M. (2009). The Analysis of Rural Poverty in Ethiopia: Regarding the three Measurements of Poverty, MA Thesis, Uppsala University, Sweden.

Tassew Woldehanna (2008). Correlates of Poverty in Urban and Rural Ethiopia. *Ethiopian Journal of Agricultural Economics*, **7** (1): 49-81.

UNDP (2005). Human Development Report 2005: International Cooperation at Crossroads: Aid, Trade and Security in an Unequal World. United Nations Development Program.

World Bank (2000). "World Development Report 2000/01: Attacking Poverty", New York: Oxford University Press.

World Development Online Indicators (2008). (www.dfid.gov.uk.).