ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

COLLEGE OF NATURAL SCIENCES

DEPARTMENT OF STATISTICS

STUNTING STATUS OF UNDER-FIVE CHILDREN IN RURAL ETHIOPIA: MULTILEVEL LOGISTIC REGRESSION ANALYSIS

BY

TIZAZU BAYKO

A Thesis Submitted to the Graduate Programs of Addis Ababa University in Partial Fulfillment of the Requirement for the Degree of Master Science in Statistics

May 2014

Addis Ababa

Ethiopia

Addis Ababa University

School of Graduate Studies

This is to certify that the thesis prepared by Tizazu Bayko, entitled: Stunting status of under-five children in rural Ethiopia and submitted in partial fulfillment of the requirements for the Degree master of science in Statistics compiles with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the examining committee:

Examiner___________________________ signature__________ Date__________

Examiner___________________________ signature__________ Date__________

Advisor___________________________ signature __________ Date__________

____

Chair of Department or Graduate Program coordinator

DECLARATION

I the undersigned, declare that the thesis is my original work, has not been presented for degree in any other University and that all sources of material used for the thesis have been duly acknowledged.

Declare by:

Name ____

Signature ______________________________________

Date ____

Place of submission: Department of statistics, college of natural science, Addis Ababa University. This thesis has been submitted for examination with my approval as a university advisor.

Name Dr. Shibru Temesgen

Signature _____________________________________

Date ____

ABSTRACT

Stunting status of under-five children in rural Ethiopia

Tizazu Bayko

Addis Ababa University, 2014

Stunting is a well-established child health indicator for chronic malnutrition related to environmental and socio-economic circumstances (WHO, 1995 and 1996). According to the 2011 EDHS report by the Ethiopia CSA, nationally, 44 percent of children under age five are stunted, and 21 percent of children are severely stunted. This study aimed to identify socio-economic and demographic determinants of under-five child stunting in rural Ethiopia using EDHS 2011 data. In order to achieve our objective descriptive, cross-tabulation, binary logistic regression and multilevel logistic regression statistical techniques were used for data analysis using socio-economic, demographic and health and environmental variables as explanatory variables, and status of stunting as the response variable. The results of the analysis show that age of child, mother's educational level, wealth index, employment status of mother, source of drinking water, types of toilet facilities, had diarrhea two weeks before the survey, had fever two weeks before the survey, educational level of partner and geographical region were found to be significant determinants of stunting status of under-five children in rural Ethiopia. The result also suggested that children from poor families were more likely to be stunted than children from rich families in rural Ethiopia. It is found that children from uneducated mothers are at higher risk of stunting. It is recommended that design and implement primary health care and nutrition programs which would fit the features of each region to safeguard children from nutritional deficiency.

ACKNOWLEDGEMENT

First, and foremost, I thank God for giving me the opportunity to pursue my graduate study at the Department of Statistics, Addis Ababa University.

I owe the deepest gratitude to Shibru Temesgen (PhD), my thesis advisor for his invaluable and constructive comments and suggestions throughout my study. Without his professional assistance and guidance this study would not be realized.

My special thanks also go to all staff members of the Department of Statistics and to my colleagues for all the encouragements and support that they provided me during my study years.

I would also like to thank my employer, the Ministry of Education, for the sponsorship that allowed me to pursue my postgraduate education.

Finally and most importantly, my heart-felt thanks go to all members of my family for all the encouragement and support throughout my study.

ACRONYMS

ACC- Administration Committee on Coordination

SCN- Sub-Committee on Nutrition

AIC - Akaike Information Criteria

BIC - Bayesian Information Criteria

Cons-constant

Chi2-chi-square

CSA- Central Statistical Agency

df- degree of freedom

DHS - demographic and health survey

EAs - Enumeration areas

EDHS- Ethiopia demographic and health survey

H//A- Height-for-Age

LL- Log Likelihood

LR- Likelihood Ratio

MDGs - Millennium Development Goals

ML- Maximum Likelihood

MOH- Ministry of health

NCHS- National Center for Health Statistics

OR - Odds ratio

PEU- Protein/Energy Under-nutrition

SAS- Statistical analysis system

SNNP-Southern Nations Nationalities and Peoples

SPSS- Statistical Package for Social Science

UN - United Nations

UNICEF- United Nation Children's Fund

W//A- Weight-for-Age

W//H- Weight-for-Height

WHO-World Health Organization

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1. Background of the study

The nutritional status of under five children is one of the indicators of household wellbeing and one of the determinants of child survival. Poor nutrition leads to ill health and ill health causes further deterioration of nutritional status. Health and nutrition are important and critical components for developing countries. In addition they are channels through which productivity and distribution goals of developing societies may be pursued effectively. As often hypothesized, the productivity of low-income persons in work and in human capital information is positively affected by health and nutritional status (Smith and Haddad, 2000).

Nutrition and health are important dimensions of human well-being. Malnutrition represents the main health problem in developing countries. Though the word “malnutrition” is associated with both under nutrition and over nutrition (Smith and Haddad, 2000). Malnutrition among women is likely to have a major impact on their own health and on the health of their children. Besides her own health, a mother’s nutritional status affects her capacity to successfully care for her children (Abey, 1995). Women in the reproductive age group (15-49 years) and children are most vulnerable to malnutrition due to low dietary intakes, inequitable distribution of food within the household, improper food storage and preparation, dietary taboos and infectious diseases. Particularly for women, the high nutritional costs of pregnancy also contribute significantly to their poor nutritional status (Woldemariam and Timotiwos, 2002).

Woldemariam and Timotiows (2002) performed a comparative study for urban, rural and combined urban and rural children. They identified region of residence, education of mother, economic status of the household and age of the child as determinants of malnutrition and health problems among urban children. For rural children, the study showed that region of residence, education of mother, education of partner, age, birth order and preceding birth interval of a child as important predictors of nutrition/health status. The combined urban and rural (national) sample results indicated that region of residence, education of mother, education of father, economic status of the household, age, birth order and birth interval of the child were found to be determinants of child nutrition and health status.

The nutritional status of infants and children under five years of age is of particular concern since the early years of life are crucial for optimal growth and development. Their nutritional wellbeing reflects household, community and national investments in family health thereby contributing both directly and indirectly to the overall development of the country. It is estimated that globally among the under 5 children 226 million are stunted, 67 million are wasted and 183 million weigh less than they should for their age (Woldemariam and Timotiwos, 2002). Since children are economic assets to the world and their future development outcome can be influenced by their nutritional status, the mechanism and consequences of malnutrition need to be understood better. The prevalence of stunting in children below five years in East Africa averages about 48 percent (ACC/SCN, 2000), which is the highest in the world.

Evidence also showed that the situation in Ethiopia is worse than in other East African countries. A review of the trends of the nutritional status of children in Ethiopia from 1993-1998 showed that the national rural prevalence of stunting increased from 60 percent in 1983 to 64 percent in 1992. Another national survey undertaken in 1998 with the inclusion of urban areas and children in the age group 3-5 months showed a relative decline in the proportion of stunted children to 52 percent (Zewditu et al., 2001). Therefore, there is a need to identify the factors that affect the nutritional status of children so that interventions can be planned to children achieve optimum growth and development.

Stunting is a well-established child health indicator of chronic malnutrition related to environmental and socio-economic circumstances (WHO, 1995 and 1996). Childhood is a period of an active growth, covering the major transformations from birth to adulthood. Adequate nutrition is needed to ensure optimum growth and development of children. Normal growth is dependent on adequate nutrition, and human body can use carbohydrate, protein and fat as a source of energy. Inadequate intake of energy may lead to malnutrition in the long run (Mamoun et al, 2005). A stunted child is a child below his or her appropriate height for age (Rice et al, 2000). Stunting affects the physical and mental outcome of the children adversely. Growth imposes a high metabolic demand throughout babyhood, particularly through the first year of life. High demands related to the increased rates of growth, higher amounts of active tissues per unit of body mass (e.g. brain), and happening of disease are among the factors making babies a vulnerable group (Mamoun et al, 2005).

According to the Ethiopia CSA report in 2011 EDHS nationally, 44 percent of children under age five are stunted, and 21 percent of children are severely stunted. In general, the prevalence of stunting increases as the age of a child increases, with the highest prevalence of chronic malnutrition found in children age 24-35 months (57 percent) and lowest in children under age six months (10 percent). Male children are slightly more likely to be stunted than female children (46 percent and 43 percent, respectively). With the exception of first order births, there is an inverse relationship between the length of the preceding birth interval and the proportion of children who are stunted. The longer the interval, the less likely it is that the child will be stunted. In this study, we will be focusing on stunting as an indicator of malnutrition.

Reducing malnutrition among children under the age of five remains a huge challenge in developing countries of the World. An estimated 230 million under-five children are believed to be chronically malnourished in developing countries (Van de Poel et. al., 2008). Similarly, about 54% of deaths among children of this age group are believed to be associated with malnutrition in developing countries (FAO, 2008). In Sub-Saharan Africa, 41% of under-five children are malnourished and deaths from malnutrition are increasing on daily basis in the region (FAO, 2008). Malnutrition is widespread in Nigeria, especially in the rural areas. This is partly due to inadequate food and nutrient supply. The 2003 Nigeria Demographic and Health Survey revealed that 38% of under-five children in Nigeria are stunted, 29% underweight and 9.2% wasted (Ajieroh, 2010). The 2004 Food Consumption and Nutrition Survey reported similar trends with 42% stunted, 25% underweight and 9% wasted (Ajieroh, 2010). These surveys indicated significant variation between the rural and urban areas with children from rural areas worse affected by malnutrition.

1.2. Statement of the Problem

The nutritional status of children in the rural area of Ethiopia is among the worst in the world. Almost one in every ten babies born in Ethiopia (97 per 1000) does not survive to its first birthday, and one in every six children dies before its fifth birthday. The WMS (2004) report shows that rural children are reported as more prone to all kind of malnutrition. the National Nutrition Strategy document, the right to the determinants of adequate nutrition is also upheld in the constitution of Ethiopia, entrusting the government to take appropriate measures to ensure that these rights are adequately protected, especially among the most vulnerable. It is said that, therefore, in fulfilling its obligations, the government has recently formulated the national nutrition strategy (FDRE, 2005).

High malnutrition rates in Ethiopia pose a significant obstacle to achieving better child health outcomes. Malnutrition among under-five children is a chronic problem in developing countries like Ethiopia. Worldwide, ten and a half million children of age under-five die every year, with 98% of these deaths reported to occur in developing countries (UNICEF, 2007). In recognition of the burden of malnutrition among under-five children, four of the eight United Nations Millennium Development Goals (MDGs) are specifically directed towards improving child health outcomes in developing countries. In particular, a reduction in the mortality of children is a key MDG, and a reduction in malnourishment among children is an important indicator of progress towards that goal.

Based on the 2011 Ethiopian Demographic and Health Survey (EDHS), children under age five, the prevalence of stunting (their z-score less than -2) at country level is reported 44 percent. The figure shows the extent to which how much of the country's potential work force is faced with growth retardation.

Even though the problem of child malnutrition in Ethiopia has been sufficiently documented, the reasons behind it are still poorly understood. There is also inconsistency across studies regarding the determinant factors behind childhood stunting. Therefore, this study attempts to investigate the major socio-economic, demographic, health and environmental determinants of stunting in Ethiopia.

1.3. Objective of the Study

The general objective of this study is to determine the prevalence of stunting and identify the various possible factors associated with stunting among under-five children in rural part of Ethiopia.

The specific objectives of this study are:

- To identify the most important socio-economic, demographic, and environmental factors associated with stunting among under-five children in rural Ethiopia.
- Examine the extent of the variation in stunting of under-five children within and between regions of the rural parts of Ethiopia.

1.4. Significance of the study

The study was conducted to investigate the main risk factors related to stunting status of under-five children. Through this research an analysis of children nutritional status in relation to demographic, socio-economic, and feeding practice is presented. It is hoped that the finding of this study will identify the risk factors for stunting among children aged less than five years in rural Ethiopia.

In addition to this, the findings could be helpful for policy making, monitoring and evaluation activities of the government and different concerned agencies. Since the study will attempt to reveal the major factors for childhood stunting, it will help to guide the end user governmental and non-governmental organizations to develop nutrition programs and set appropriate plans to tackle the existing health and nutrition problems. Finally, the study could be used as a stepping stone for further studies.

CHAPTER TWO

LITERATURE REVIEW

2.1. Concepts and Definition

Adequate, appropriate and safe food and nutrition means the regular and sufficient consumption of nutritious foods across the life span, including breastfeeding, to support normal growth and development of children and promote physical, emotional, and social well-being for all people.

The word "anthropometric" is generally meant to represent the measure of people's growth indicators such as weights and heights (related to their age and sex). It is used for growth assessment and is a single measurement that best defines the nutritional status of a child (Blossner and De Onis, 2005). According to this measure, the nutritional/health status of children is determined by comparing growth indicators with the distribution of the same indicators for "healthy" (as reference group), and identifying "extreme" or "abnormal" departures from this distribution. The international reference standard that is most commonly used (and recommended by the WHO) is that of the data on the weights and heights of a statistically valid population (US National Center for Health Statistics (NCHS)) of healthy infants and children in the US.

2.1.1. Malnutrition

Adequate nutrition is an integral part of health, happiness, independence, quality of life, and physical and mental functioning. In 1974, C.E. Butterworth suggested in Nutrition Today that malnutrition was the skeleton in the hospital closet: he was convinced that malnutrition occurring in acute care hospitals was "physician-induced". At that time, studies had yet to document that an equally ominous skeleton lurked in the nursing home closet. Since then, studies using a variety of measurements and sampling frames have shown that 35 to 85 percent of nursing home residents are malnourished. This encompasses a wide array of conditions—obesity, lipid disorders, protein/energy under-nutrition (PEU), and vitamin and mineral deficiencies. All except obesity are prevalent among nursing home residents.

In developing countries, stunted growth is a common problem affecting a large percentage of children, and it is one of the main health problems in Ethiopia. Malnutrition of young children occurs mainly under conditions of extreme poverty and deprivation; it has a huge impact on their linear growth, activity, and level of their cleverness and intelligence. Stunting is also closely linked to impairments in mental development, and it is associated with high morbidity and mortality rates among children (Rolfes et al, 2004).

2.1.2. Definition of Stunting

Stunting as an expression was introduced in the 1970s by J.C.Waterlow describing the linear growth retardation in children that results in decreasing their height with referring to age caused by prolonged nutritional destitution (Kikafunda et al, 1998).

Stunted growth is a reduced growth rate in human development. Once established, stunting and its effects typically become permanent. Stunted children may never regain the height lost as a result of stunting, and most children will never gain the corresponding body weight. It also leads to premature death later in life (Olivieri et al, 2008). Stunting, or chronic malnutrition, defined on the basis of the height to age ratio, shows malnutrition resulting from cumulative inadequacies in the child's nutritional status (Shaikh et al, 2003).

According to World Health Organization recommendations, the cuts off put to define stunting were less than two z-scores of height for age. Stunting has been considered as an indication of chronic malnutrition or something that has happened in the past, but this is a misconception since although the nutritional insult may have happened in the past the consequent process is still ongoing. Stunted children are more likely to get sick and die than underweight or wasted children are (Shrimpton et al, 2003).

Stunting as a well being health indicator is used to describe the level of child malnutrition. Stunting or low height-for-age measures linear growth retardation and cumulative growth deficit. It also indicates the effect of past nutritional insult in the life of the child (Genebo and Girma, 2002). The proportion of children who are more than 2 standard deviation units (z-score) below the median of the international reference population are considered moderately stunted and those who fall more than 3 SD units below are severely stunted (Padmadas, Hutter and Willekens, 2002).

Stunting is a cumulative process of poor growth that primarily occurs before the age of 3 years and is not easily reversed, whereas child growth is considered a good indicator of overall socioeconomic development and human welfare in developing countries. Growth retardation is a physical indicator of a broad spectrum of nutritional deficiencies and is often linked to poor mental development (Lang, 1998). Growth stunting is defined by comparing measurements of children's heights to the NCHS growth reference population; children who fall below the fifth percentile of the reference population in height-for-age are defined as stunted, regardless of the reason of their shortness (Lewit and Kerrebrock, 1997).

2.2. Prevalence of stunting

Data on the prevalence of protein energy malnutrition in developing countries indicate that on average, stunting affects over 40% of fewer than 5 years children, and stunting rates among children are highest in Africa and Asia. Thin and short were used in the past to refer to children with abnormal health and nutritional status. Nowadays; wasting and stunting are believed to insinuate deviation behind the range of thinness and shortness that might be regarded as normal in some cases. It soon became apparent that stunting is generally more common than wasting, and in some populations, particularly in Asia, over 50 % of children could be classified as stunted (De Onis et al, 1993).

In a study drawn on the experience of 63 developing countries over the 25-years period to identify the determinants of child malnutrition for each developing region, only seven countries have a higher prevalence of child stunting than Bangladesh. While Bangladesh has the highest prevalence of childhood underweight among all countries in the world, except North Korea, the percentage of children aged less than 5 years with stunting decreased from 64.2% in 1992 to 48.3% in 2000 and 42.4% in 2005 (Rahman, Mostofa and Nasrin, 2009).

In developing countries, stunting affects about one-third of the children below five years of age. The highest levels of stunting are observed in eastern Africa, where on average 48.1 % of under five children are stunted. In 1998 a demographic health survey (DHS) done in Kenya showed that the prevalence of stunting was 22%. Another DHS done in 2003 showed little improvements with the prevalence of stunting dropped to 20%. On the other hand in the central province, lower improvements were observed with the prevalence of stunting reduced to 28%. By the year 2007 some11% of males and 16% of females were found to be stunted in Kenya (Veronicah, Kogi-Makau, and Muroki, 2007).

Evidence also showed that the situation in Ethiopia is worse than in other East African countries. A review of the trends of the nutritional status of Ethiopian children from 1983-1998 showed that the national rural prevalence of stunting increases from 60 percent in 1983 to 64 percent in 1992. Another national survey undertaken in 1998 with the inclusion of urban areas and children in the age group 3-5 months showed a relative decline in the proportion of stunted children to 52 percent (Zewditu et al.,2001). A few local studies (Getaneh et al., 1998; Genebo et al., 1999; Gugsa Yimer, 2000) on child nutrition have also shown similar results (a more than 40 percent prevalence in stunting) and confirmed that malnutrition, i.e., stunting, is one of the most important public health problems in this country.

2.3. Nutritional Assessment of Stunting

Globally, height-for-age has been used as the indication of growth assessment according to WHO and NCHS standards. These international standards test the accuracy of anthropometric measures from around the world. Anthropometric measurements are converted into three indexes: H//A, W//A, and W//H which have been expressed as Z score relative to the international (NCHS, WHO, and CDC) reference population to standardize the distribution. A child's nutritional status is then categorized by his or her height-for-age or weight for height Z scores. A child is categorized as stunted if his or her height is less than that of the child of the same age with a value 2 standard deviations below the reference median height for age (Ricci and Becker, 1996).

For children less than 2 years of age, recumbent length was used in the anthropometric calculations, and for children of age 2 years and higher, standing heights were used. Low anthropometry was defined as a Z-score < 2 SD. (Hassan et al, 1997). The stunting is considered mild if ≤ the mean (- 1SD), moderate if ≤ the mean (- 2SD), and severe if ≤ the mean (- 3SD).

2.4. Socioeconomic characteristics

Socioeconomic characteristics such as household economic status, maternal socioeconomic characteristics (mothers' education, mothers' employment status and their household status relative to men)**,** household size, and likes are important in child health outcomes.

2.4.1. Household economic welfare

According to Christiaensen and Alderman (2004), sustained income growth of 2.5 percent per adult over a fifteen-year period could be associated with a 3-6 percent decline in chronic child malnutrition in Ethiopia. However, they argued that income growth alone might not be sufficient to alleviate child malnutrition. Similarly, Glewwe *et al.* (2002) observed that, due to rapid economic growth in Vietnam since 1986, a dramatic decline in poverty and child stunting, wasting and underweight was achieved. Based on their empirical investigation, they concluded that growth in household income, although not very large, had a positive impact on child nutrition in Vietnam during the 1990s. They also noted that, over time, child stunting, wasting and underweight declined within each quintile even after adjustment for change in income was made, which suggests that there are other factors, in addition to income growth, which led to improvements in child nutrition.

The better off households have better access to food and higher cash incomes than poor households, allowing them a quality diet, better access to medical care and more money to spend on essential non-food items such as schooling, clothing and hygiene products (Stalin p. et al., 2013). The study showed that the prevalence of stunting among under-five children who belongs to class poorest and class poor socio-economic status was 47.2% and 63.4% respectively. Under five children belonging to class medium socio-economic status were stunted (46.7%).

2.4.2. Household Factor

A growing number of studies recognize the important role that household composition and parental levels of education play in shaping children's nutritional outcomes. Christiaensen and Alderman (2004) found that larger family size results in stunting among children in Ethiopia. They argue that economies of scale in time for childcare and expenditure can be enjoyed in large families and that children benefit from parent's accumulated experiences in care of young children. Various studies have concluded that parental education, especially mother's education, is a key element in improving children's nutritional status (Moen, 1993; Christiaensen and Alderman, 2004). The implication is that schooling is associated with child nutrition only if it can improve mother's nutritional knowledge. An educated mother is likely to have a higher income (which can directly affect her children's health and nutrition) and power in the household and community which will put her in a better position to make decisions about her children's needs (Moen, 1993). The literature also notes that, particularly where the general level of education of the community is low, the level of education of female and male members of the household could be particularly important in indirectly influencing child nutritional status (Basu and Foster, 1998; Gibson, 2001). More specifically, Christiaensen and Alderman (2004) found that the effect of maternal education is about twice as important as that of paternal education.

2.5. Demographic characteristics

Studies show that while the main causes of malnutrition appear to change with age of children, in most cases older children are found to be associated with increased malnutrition. For example, in Ethiopia report on health and poverty by WB and MOH (2005) using descriptive statistical method show that older children have a higher likelihood of being underweight and stunted relative to children who are less than a year old. At country level, all the four welfare monitoring surveys from 1996-2004 using descriptive statistics have revealed that boys are more vulnerable to malnutrition than girls with respect to the three indices (wasting, stunting, and underweight). Similar results are also reported from some case studies and official surveys (Sentayehu, 1994; Christiaensen and Alderman, 2001; Alemu et al. 2005b). Various reasons behind this gender differential are given in the literature. Alemu et al. (2005a), for example, argue that this could be due to genetic differences between male and female children and, due to girls' greater access to food through their gender-ascribed role in contributing to food preparation. However, using the 2000 Ethiopia Demographic and Health Survey data, Silva (2005) did not find the coefficient on the child's gender to be significant in any of the regressions, suggesting there is no gender bias affecting the nutritional status of children in Ethiopia. Similarly, Bilisuma (2004), using probit model, could not find sex of a child to have a significant impact on the probability of being stunted.

It is expected that parents give less attention to older children when they give birth to a new child who needs much attention and care. One study showed that stunting is rare in birth order 2-3 (Sommerfelt et al., 1994), and higher birth order (5+) is positively associated with child malnutrition (Jeyaseelan, 1997). Using Young Lives data of children between the age of 6 and 18 months in Peruvian, Alemu et al. (2005a) did not find birth order to be significantly associated with any of the three indicators considered (wasting, stunting, and underweight) for the whole sample by using logistic regression. However, the results stratified by location show that birth order is associated with wasting and underweight for urban children, while the likelihood of being wasted decreases with higher birth order in rural areas. On the other hand, unlike expectations, based on the National Rural Nutrition Survey of 1992 in Sidamo, Sentayehu (1994) by using linear regression analysis found positive sign of birth order on height-for-age and weight- for-age equation. The sign is unexpected because high birth order is expected to adversely affect the quantity and quality of resources that could be allocated to the children in the household.

2.6. Health and Environmental Characteristics

Access to unsafe water and unsanitary disposal of wastes are regarded as the main causes of infectious diseases such as diarrhea and intestinal parasites (UNICEF cited in Smith et al, 2005). Where there is a better access to safe water and quality sanitation, the incidence of various illnesses will decline (Smith and Haddad, 2000). World Bank (2006) stated that improving access and quality of safe water not only reduces transmission of waterborne diseases but also saves women the extra time they spend on carrying water which can be allotted to child care and feeding or income generating activities.

Silva (2005) examined the impact of access to basic environmental services, such as water and sanitation, on children’s nutritional status using probit analysis and these are found to be insignificant. However, Silva noted that the results for the model including community environmental sample indicate the coefficients on the proportion of households with access to these services are highly significant in the underweight equations suggesting a spillover effect of other household's access to these services.

Christiaensen and Alderman (2001) found possession of a tap and a flush toilet to have a positive effect on child height. However, access to other sources of drinking water which are generally deemed safe such as public taps and protected wells were not found to positively affect children's height. Similar results for underweight are also found by Woldemariam and Timotiwos, 2002; Alemu et al, 2005a.

Considering usage of a tap and a flush toilet, contrary results to Christiaensen and Alderman (2001) are reported in Alemu et al. (2005a) and Alemu et al (2005b) for the case of wasting in urban and rural areas, respectively. The reason suggested for the former is that it could be because of the unhealthy conditions of communal latrines in slum areas, while the rural case is assumed that people may still prefer to use the open field rather than unfamiliar pit latrines. Many studies use the distance between the household's home and the nearest health facility to proxy for access to health care (see Christiaensen and Alderman, 2001; Alemu et al, 2005a, Alemu et al, 2005b). In addition to these studies, however, Collier et al., 2002 and Abay Asfaw (1995) argue that usage of health services is sensitive not just to the distance to the nearest facility but also the quality (such as availability of material inputs and drugs, the number and qualification of staff, user fees, etc) of care provided.

A household's access to facilities is likely to be correlated with community characteristics. Households living in wealthier communities might have a relatively healthy environment, which implies better sanitation facilities, access to clean water and healthcare facilities (Glewwe *et al.*, 2002). Water and sanitation play a particularly important role in child nutrition due to their impact on diarrheal diseases.

2.7. Review on Statistical Models

Binary logistic regression is typically used when the dependent variable is dichotomous and the independent variables are either continuous or categorical variables. When the dependent variable is not dichotomous and is comprised of more than two cases, a multinomial logistic regression can be employed. Logistic regression is a popular modeling approach when the dependent variable is dichotomous or polytomous. Hosmer and Lemeshow (2000) have described logistic regression focusing on its theoretical and applied aspect. Logistic model, as compared to its competitor, the probit model, is less sensitive to outliers and easy to correct a bias (Copas, 1988).

Often the outcome variable in social data is, in general not continuous, instead is binary. In such a case, binary logistic regression is a useful way of describing the relationship between one or more independent variables and a binary outcome variable that has only two possible values. Indeed, a generalized linear model is used for binary logistic regression. The most attractive feature of a logistic regression model is that it neither assumes linearity in the relationship between the covariates and the outcome variable, nor does it require normally distributed variables. It also does not assume homoscedasticity and in general has less stringent requirements than linear regression models. Thus logistic regression is used in a wide range of applications leading to binary dependent data analysis (Agresti, 2002).

Multilevel analysis is a methodology for the analysis of data with complex patterns of variability, with a focus on nested sources of variability. When the data structure is hierarchical with elementary units at level 1 nested in clusters at level 2, which in turn may be nested in (super) clusters at level 3, and so on. The latent variables, or random effects, are interpreted as unobserved heterogeneity at the different levels which induce dependence among all lower-level units belonging to a higher-level unit. Random intercepts represent heterogeneity between clusters in the overall response and random coefficients represent heterogeneity in the relationship between the response and explanatory variables. (Goldstein, 2003).

The multilevel logistic regression analyses consider the variations due to hierarchy structure in the data. It allows the simultaneous examination of the effects of group level (cluster and division) and individual level variables on individual level outcomes while accounting for the non independence of observations within groups. Also this analysis allows the examination of both between group and within group variability as well as how group level and individual level variables are related to variability at both levels (Goldestein and Rasbash, 1996).

CHAPTER THREE

DATA AND METHDOLOGY

3.1. Data Source

The source of data for this study was the 2011 Ethiopia Demographic and Health Survey (EDHS) which is obtained from Central Statistical Agency (CSA). It is the third major survey designed to provide estimates for the health and demographic variables of interest for the following domains: Ethiopia as a whole, urban and rural areas of Ethiopia (each as a separate domain), and all geographic areas (regions). In the 2011 EDHS a representative sample of approximately 17,018 households from 624 clusters was selected. The sample was selected in two stages. In the first stage, 624 clusters (187 urban and 437 rural) were selected from the list of Enumeration Areas (EA). In this study we try to identify the health and demographic variables that affect stunting of under-five children in the rural Ethiopia.

3.2. Variables of the Study

As discussed in the literature review, socio-economic, demographic, health and environmental characteristics are considered as the most important determinants of stunting in under-five children.

3.2.1. The Response Variable

In different studies the response variable height-for-age (stunting) measurement status was expressed in Standard Deviation (SD) units (Z-score) from the median of the reference population. Children with a measurement of <-2 SD units from the median of the reference population were considered short for their age (stunted) and children with measurement of (≥ -2SD) units from the median of the reference population were not stunted.

In this study, the measurement stunting status of under-five children was calculated using the new Child Growth Standards released by the World Health Organization in April 2006. The new Standards are the result of an intensive study initiated by WHO in 1997 to develop a new international standard for assessing the physical growth nutritional status and motor development in all children from birth to age five. WHO and its principal partner, the United Nations University, undertook a Multicentre Growth Reference Study which is a community-based multicounty project involving more than eight thousand children from Brazil, Ghana, India, Norway, Oman, and the United States of America. The measures are presented with two implied decimal places.

To determine the level of stunting, the dependent variable was expressed as a dichotomous variable category 0 if not stunted (≥-2SD) and category 1 if stunted (<-2SD). In view of this, the response variable, the status of stunting of the i^{th} child was measured as a dichotomous variable:

$$Y_i = \begin{cases} 0, not\ stunted\ if\ Z - score \geq -2SD\ \ from\ the\ median\ of\ reference \\ 1, stunted\ if\ Z - score < \ -2SD\ \ from\ the\ median\ of\ reference \end{cases}$$

Y_i = represent the stunting status of i^{th} child.

3.2.2. Explanatory Variable

The predictor variables to be studied as determinants of under-five children stunting are grouped into socio-economic, demographic, and health and environmental factors.

Table 3.1 Description of Variables and Coding

The response variable

Variable	Representation of Variables	Categories
Stunting	Y	0, Not Stunted 1, Stunted

Explanatory Variables

1. Demographic Variables

Variables	Categories
Sex of Child (SC)	0 = Female 1 = Male
Age of Child (AC)	0 = 0-5 months 1 = 6-11 months 2 = 12-23 months 3 = 24-35 months 4 = 36-47 months 5 = 48-59 months
Birth order of Child (BOC)	0 = 1 1 = 2-3 2 = 4-5 3 =6+

2. Socio-Economic Variables

Variables	Factor categories
Mother's education (ME)	0 = no education 1 = primary education 2 = secondary and above
Employment status of mother (ESM)	0 = unemployed 1 = employed
Wealth index (WI)	0 = poor 1 = medium 2 = rich

Education of husband/partner (ES)	0 = no education 1 = primary education 2 = secondary and above
Number of household member (NHHM)	0 = 1-4 1 = 5-9 2= 10 and above
Region (RGN)	0 = Tigray 1 = Afar 2 = Amhara 3 = Oromia 4 = Somali 5 = Ben-gumuz 6 = SNNP 7 = Gambela 8 = Harari 9 =Dire Dawa

3. Health and Environmental Characteristics

Variables	Factor categories
Source of drinking water (SDW)	0 = non-improved source (surface water, tanker truck, unprotected well/spring, other's) 1 = improved source (piped into dwelling, piped to yard/plot, public tap, protected well/spring, rain water, bottled water)
Had diarrhea in the two weeks before survey (DIARRRHEA)	0 = no 1 = yes
Had fever in the two weeks before survey (FEVER)	0 = no 1 = yes

Types of toilet facilities	0 = no facilities (bush/field toilet, hanging toilet, others) 1 = have facilities (pit toilet, flush/pour toilet)

3.3. Methodology

3.3.1. The Binary Logistic Regression Model

Binary logistic regression analysis examines the influence of various factors on a dichotomous outcome by estimating the probability of the event's occurrence. It does this by examining the relationship between one or more independent variables and the log odds of the dichotomous outcome by calculating changes in the log odds of the dependent as opposed to the dependent variable itself. The log odds ratio is the ratio of two odds and it is a summary measure of the relationship between two variables. The use of the log odds ratio in logistic regression provides a more simplistic description of the probabilistic relationship of the variables and the outcome in comparison to a linear regression by which linear relationships and more rich information can be drawn.

Regression methods have become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables. The outcome variable in binary logistic regression is binary or dichotomous while the outcome variable in linear regression is continuous. On the other hand the error term in linear regression is normally distributed with mean zero and some variance that is constant across levels of the independent variables while the error term in binary logistic regression is distributed binomially, not normally. Consequently the response variable in logistic regression is not restricted to normality in case of parameter estimation. Because of this logistic regression is the most popular method of analyzing binary response data. There are two primary reasons for choosing the logistic regression. First from mathematical point of view it is an extremely flexible and easily used function and second it lends itself to a clinically meaning full interpretation (Hosmer and Lemeshow, 2000).

Logistic regression is used regularly rather than discriminate analysis when there are two categories of the dependent variable. Discriminate analysis strictly requires the continuous independent variables (though dummy variables can be used as in multiple regressions).Thus, in instances where the independent variables are categorical, or a mix of continuous and categorical, and the dependent variable is categorical, logistic regression is appropriate (Agresti, 2007).

Logistic regression determines the impact of multiple independent variables presented simultaneously to predict membership of one or other of the two dependent variable categories. There are two main uses of logistic regression: the first is the prediction of group membership. Since logistic regression calculates the probability of success over the probability of failure, the results of the analysis are in the form of an odds ratio. Logistic regression also provides knowledge of the relationships and strengths among the variables.

The assumptions required in logistic regression are less restrictive than those for ordinary least squares regression. There is no formal requirement for multivariate normality, homoscedasticity, or linearity of the independent variables within each category of the response variable. However, the assumptions that apply to logistic regression model include: meaningful coding, inclusion of all relevant and exclusion of all irrelevant variables in the regression model and low error in the explanatory variables.

Model: Let Y_{nx1} be a dichotomous outcome random vector with categories 1 (a child experienced under-five stunting) and 0 (a child didn't experience under-five stunting). Let X, an n×(k+1) matrix denotes the collection of k-predictor variables and, β be a (k+1) ×1 vector of parameters, then the data layout of explanatory variables is given by: $X = (1, x_1, x_2, \dots, x_k)'$ $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)'$

The first (leading) column is a column of 1's. It corresponds to the constant or intercept of logistic regression equation. X without the leading column of 1s, is termed as predictor data matrix. Then, the conditional probability of Y given X is given by:

$$P(Y = 1|X) = \frac{e^{g(x)}}{1+e^{g(x)}} = \frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_k X_k}}{1+e^{\beta_0+\beta_1 X_1+\cdots+\beta_k X_k}} = \frac{\exp(\boldsymbol{X}'\boldsymbol{\beta})}{1+\exp(\boldsymbol{X}'\boldsymbol{\beta})} \qquad (3.1)$$

From (3.1) the odds of success is given by:

$$Odds(Y = 1) = \frac{p}{1-p} = \exp\left(\boldsymbol{X}'\boldsymbol{\beta}\right) \quad (3.2)$$

In logistic regression analysis, it is assumed that the explanatory variables affect the response variable through a suitable transformation of odds of the probability of success. This transformation is a suitable link function of p and is called the logit - link, which is defined as:

$$logit(\pi) = log\left(\frac{P}{1-P}\right) = log(\exp\left(\boldsymbol{X}'\boldsymbol{\beta}\right)) \quad (3.3)$$

The transformed variable, denoted by logit (π) is the log-odds and is related to the explanatory variables as:

$$logit(\pi) = g(x) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k = \boldsymbol{X}'\boldsymbol{\beta} \quad (3.4)$$

The above equations give suitable representations of the success probability, odds, and log-odds. Indeed, these representations facilitate interpretations of parameter estimates. The parameter β_i refers to the effect of X_i on the log- odds for Y = 1 controlling the other X's in the model. For predictor variables having L levels ($L \geq 2$), interpretation can be made by making one of the L-levels as a reference category.

Odds ratio: for a probability of success p, the odds of success are defined to be the ratio of probability of success p (in our case the probability of under-five child stunting or y=1) to probability of failure 1- p (in our case the probability of under-five child was not stunted or y=0).

$$Odds = \frac{p}{1-p} \quad (3.5)$$

The odds are non-negative with value greater than one when a success is more likely than a failure and between 0 and 1 when the appropriate is true. The odds ratio denoted as OR is defined as the ratio of the odds for y=1 to the odds for y=0 and is given by the equation:

$$OR = \frac{p(1)}{1-p(1)} * \frac{1-p(0)}{P(0)} \quad (3.6)$$

The odds ratio is a measure of association which has found wide use, as it approximates how much more likely (or unlikely) it is for the outcome to be present among those with y=1 than among those with y=0.

3.3.1.1 Parameter Estimation

The logistic regression model just developed is a generalized linear model with binomial errors and logit link. Instead of finding the best fitting line by minimizing the sum of the squared residuals, as we did with OLS regression, we use Maximum Likelihood (ML) estimation to obtain estimates for the coefficients in the logit equation given in (3.4).

ML methods seek to maximize the log likelihood, LL, which reflects how likely the observed values of the outcome may be predicted from the observed values of the predictors. ML method for this case is based on Newton-Raphson iteratively reweighted least square algorithm.

Suppose (y_1, y_2,…,y_n) represent the n dependent random observations corresponding to the random variables (Y_1, Y_2,…,Y_n). Since the Y_i is a Bernoulli random variable, the probability function of Y_i is $f(y_i) = \pi_i^{y_i}(1-\pi_i)^{1-y_i}$; y_i =0 or 1 i=1, 2, 3… n. Since Y's are assumed to be independent, the joint probability function or likelihood function is given by:

$$L = l(y;\beta) = \prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{1-y_i}$$

L is the likelihood function of the parameters. The maximum likelihood principle states that the estimates of the parameters maximize the likelihood function L(β). However, we take the log of the likelihood function which is called log likelihood and differentiate it with respect to the β's to get the estimates.

$$Loglikelihood = LL = logL(\beta) = \sum_{i=1}^{n} y_i(\beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k) - \sum_{i=1}^{n} log\{1 + \exp(\beta_0) + \beta_1 X_1 + \cdots + \beta_k X_k\} - \sum_{i=1}^{n} y_i \qquad 3.7$$

Differentiating (3.7) with respect to β and equating to zero, we will get the estimates of β. Manipulating this will be tiresome because of its complexity. But there are statistical soft-wares, such as SAS, STATA and MLWIN, available to solve this iteratively based on Newton- Raphson iterative method.

3.3.1.2. Test of Goodness of Fit

Once a model has been developed, we would like to know how effective the model is in predicting the outcome variable. This is referred to as goodness-of-fit. In testing the hypothesis that the model fits the data, the two common approaches are Pearson's X^2 and the likelihood ratio tests; (Agresti 2002).

The Hosmer-Lemeshow test is another alternative to see adequacy of the model. In this approach, data are divided into g groups (usually g=10). From each group, the observed and expected number of events will be computed. Then, the Hosmer-Lemeshow test statistic is given by:

$$\hat{C} = \sum_{i=1}^{g} \left(\frac{(O_i - E_i)^2}{V_i} \right) \qquad 3.8$$

Where $E_i = nP_i$, $V_i = nP_i (1-P_i)$, g is the number of group, O_i is observed number of events in the i^{th} group, E_i is expected number of events in the i^{th} group, and V_i is a variance correction factor for the i^{th} group. If the observed number of events differs significantly from what is expected by the model, the statistic $\hat{C}$ will be large and there will be evidence against the null hypothesis that the model is adequate to fit the data. This statistic has an approximate chi-square distribution with (g-2) degrees of freedom. If the calculated value of the Hosmer and Lemeshow goodness-of-fit test statistic is not significant, the model is adequate to fit the data at an acceptable level.

The likelihood ratio (LR) test, which is defined as $-2[Lo - L_1]$ (where L_0 and L_1 are the maximized log likelihoods under the null and alternative hypothesis, respectively) will be used to test the null hypothesis that the k-coefficients for the covariates in the model are not important in explaining the response variable against the alternative that at least one of the covariates is important. Under the null hypothesis, the LR is distributed as chi-square and if it is significant, we reject the null hypothesis and conclude that at least one of the k-covariates included in the model are important in predicting the outcome variable.

The separate effects of each predictor variable in explaining the outcome variable will be made by postulating the null hypothesis that: $H_i: \beta_i = 0$ against the alternative:$H_i: \beta_i \neq 0, i = 1,2, \ldots, n$. The appropriate test statistic is the Wald chi-square statistic given by:

$$W = \frac{\hat{\beta}_i^2}{var(\hat{\beta}_i)}$$

For large sample size this statistic has an approximate chi-square distribution with one degree of freedom. The Wald statistic, however, has some undesirable properties for large coefficients as standard error is inflated lowering the Wald statistic (chi-square) value and leading to type II errors. On the other hand, Agresti (2002) suggested that the likelihood ratio (LR) test is more reliable for small sample sizes than the Wald test. Accordingly, this study will use the Wald and/or LR test statistic to assess the significance of each predictor variable.

3.3.1.3 Model diagnostics

Diagnostics are certain quantities computed from the data with the purpose of pinpointing influential points after which these influential points can be removed or corrected. Influence statistics tell us how much some feature of the model changes when a particular observation is deleted from the model fit.

DFBETAS assess the effect of an individual observation on the estimated parameter of the fitted model. A DFBETAS diagnostic is computed for each observation for each parameter estimate. It is the standardized difference in the parameter estimate due to deleting the corresponding observation. The DFBETAS are useful in detecting observations that causes instability in the selected coefficients.

Leverage values measure the influence of a point on the fit of regression. The centered leverage ranges from 0 (no influence on the fit) to (n-1)/n or it is from 0 to 1.

Cook's distance (D_i) is a measure of how much the residual of all cases would change if a particular case were excluded from the calculation of the regression coefficients. A large Cook's distance indicates that excluding a case from computation of the regression statistics changes the coefficients substantially (Cook and Weisberg, 1982). The Cook's distance, D_i, of observation i is:

$$D_i = \frac{\sum_{j=1}^{n} \left(\hat{y}_j - \hat{y}_{j(i)}\right)^2}{p\, MSE},$$

where,

- $\hat{y}_j$ is the *j*th fitted response value.
- $\hat{y}_{j(i)}$ is the *j*th fitted response value, where the fit does not include observation *i*.
- *MSE* is the mean squared error.
- *p* is the number of coefficients in the regression model.

Cook's distance is algebraically equivalent to the following expression:

$$D_i = \frac{r_i^2}{p\,MSE}\left(\frac{h_{ii}}{(1-h_{ii})^2}\right),$$

where r_i is the i^{th} residual, and h_{ii} is the i^{th} leverage value.

3.3.2. Multilevel Analysis

The 2011 EDHS data set used for this study was based on multistage stratified cluster sampling. The structure of data in the population is hierarchical, and a sample from such a population can be viewed as a multistage sample. For multistage clustered samples, the dependence among observations often comes from several levels of the hierarchy. In order to draw appropriate inferences and conclusions from multistage stratified clustered survey data we may require tricky and complicated modeling techniques like multilevel modeling (Agresti, 2007).

In multilevel research, the structure of the data in the population is hierarchical, and a sample from such a population can be viewed as a multistage sample. First we take a sample of units from the higher level (in our case from the 9 regions and one city administrations), and next we sample the sub units from available units (in our case from children). In such samples the individual observations are generally not completely independent. Because of cost, time and efficiency considerations, stratified multistage samples are the norm for sociological and demographic surveys.

This clustering sampling scheme often introduces multilevel dependency or correlation among the observations that can have implications for model parameter estimates. For multistage clustered samples, the dependence among observations often comes from several levels of the hierarchy. The problem of dependencies between individual observations also occurs in survey research, where the sample is not taken randomly but cluster sampling from geographical areas is used instead. In this case, the use of single level statistical models is no longer valid and reasonable.

3.3.2.1 Two-Level Model

In this study, the clustering of the data points within geographical regions offers a natural 2-level hierarchical structure of the data, i.e. children are nested within regions. Let y_{ij} be the binary outcome variable, coded '0' or '1', associated with level-one unit i nested within level two unit j. Also let p_{ij}be the probability that the response variable equals 1, and $p_{ij} = pr(y_{ij} = 1)$. Here, y_{ij} follows a Bernoulli distribution. Like the logistic regression the p_{ij} is modeled using the link function, logit. The two-level logistic regression model can be written as,

$$log\left(\frac{P_{ij}}{1 - P_{ij}}\right) = \beta_0 + \beta_1 x_{ij} + u_{0j}$$

where u_{0j} is the random effect for the jth region, β_0 and β_1 are coefficients

x_{ij} = observation for i^{th} child in the j^{th} region.

Therefore, conditional on u_{oj}, the y_{ij}'s can be assumed to be independently distributed. Here, u_{oj} is a random quantity and follows $N(0, \sigma_u^2)$. The basic data structure of two level logistic regression is a collection of N groups (units at level-two (regions)) and within group j (j=1, 2,…, N) a random sample of n_j level-one units. The outcome variable is dichotomous and denoted by $y_{ij}\ (i = 1,2, \dots, n_j, j = 1,2, \dots, N\ )$ for ith child in region j.

Let the success probability in region j be denoted by p_j. The dichotomous outcome variable for the individual i in group j, y_{ij}; which is either 0 or 1 can be expressed as sum of the probability in group j, p_j (the average proportion of j levels in group j, $E(y_{ij}) = p_j$) plus some individual-dependent residual ε_{ij} , that is, $y_{ij} = p_j + \varepsilon_{ij}$ the residual term is assumed to have mean zero and variance, $var(\varepsilon_{ij}) = p_j(1 - p_j)$.

3.3.2.1.1 Testing heterogeneous proportions

For the proper application of multilevel analysis the first logical step is to test heterogeneity of proportions between the groups or regions. For this purpose we use two tests: a chi-square based nonparametric test and a parametric test. The parametric test will be discussed in the subsequent sections. In this section we present the nonparametric test. To test whether there are indeed systematic differences between the groups, the well-known chi-square test for contingency table can be used. First, we consider the chi-square test and then discuss the results we obtain on the basis of the test. The test statistics is:

$$x^2 = \sum_{j=1}^{N} n_j \, (\frac{(\hat{p}_j - \hat{p}_.)^2}{\hat{p}_.(1-\hat{p}_.)} ,$$

Where,

$\hat{P}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_{ij}$, the proportion of stunted children in region j,

$\hat{P}_. = \frac{1}{M} \sum_{j=1}^{N} \sum_{i=1}^{n_j} Y_{ij}$, the overall proportion of stunted children with $M = \sum_{j=1}^{N} n_{j.}$

This statistic x^2 chi-square statistic follows approximately central chi-square distribution with $N - 1$ degrees of freedom.

Estimation of between and within -groups variance: the theoretical variance between the group dependent probabilities, i.e., the population value of var($\hat{P}_j$) can be estimated by:

$\hat{\tau}^2 = S_{between}^2 - \frac{S_{within}^2}{\hat{n}}$, where

$\tilde{n} = \frac{1}{N-1}\left\{M - \frac{\sum_{j=1}^{N} n_j^2}{M}\right\}$, Between- groups variance: $S_{between}^2 = \frac{\hat{P}_.(1-\hat{P}_.)}{\tilde{n}(N-1)} X^2$ and Within-groups variance**:**

$$S_{wihin}^2 = \frac{1}{M-N}\sum_{j=1}^{N} n_j P_{.j}\left(1 - P_{.j}\right)$$

3.3.2.1.2 The Empty Logistic Regression Model

This is the simplest case of hierarchical two level model for a dichotomous outcome variable in which there are no explanatory variables at all. This model only contains random groups and random variation within groups. We focus on the model that specifies the transformed probabilities $f\left(p_j\right)$ to have a normal distribution. This is expressed, for a general link function $f\left(p_j\right)$, by the formula:

$$f\left(p_j\right) = \beta_0 + u_{0j}$$

where β_0 is the population average of the transformed probabilities and u_{0j} the random deviation from this average for group j. If $f\left(p_j\right)$ is the logit function, then $f\left(p_j\right)$ is just the log-odds for group j. Thus, for the logit link function, the log-odds have a normal distribution in the population of groups, which is expressed by:

$$logit\left(p_j\right) = \beta_0 + u_{0j}$$

For the deviations u_{0j} it is assumed that they are independent random variables with a normal distribution with mean zero and variance σ_0^2.

This model does not include a separate parameter for the level-one variance. This is because the level-one residual variance of the dichotomous outcome variable follows directly from the success probability, as indicated by equation $var\left(\varepsilon_{ij}\right) = p_j\left(1 - p_j\right)$. The probability corresponding to the average value β_0 denoted by π_0 is defined by $f(\pi_0) = \beta_0$.

For the logit function, the so-called logistic transformation of β_0, is defined by:

$$\pi_0 = logit(\beta_0) = \frac{exp(\beta_0)}{1 + exp(\beta_0)}$$

Note that due to the non-linear nature of the logit link function, there is no simple relation between the variances of the deviations u_{0j}. However, there is an approximate formula which is valid when the variances are small given by

$$var(p_0) = \left(\pi_0(1-\pi_0)\right)^2 \sigma_0^2$$

3.3.2.1.3 The Random Intercept with fixed effect Model

In the random intercept logistic regression model, the intercept is the only random effect meaning that the groups differ with respect to the average value of the response variable. But the relation between explanatory and response variables can differ between groups in more ways. The random intercept model expresses the log-odds (the logit of p_{ij}) as a sum of a linear function of the explanatory variables. That is,

$$logit(p_{ij}) = log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{0j} + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \dots + \beta_k x_{kij}$$

$$= \beta_{0j} + \sum_{h=1}^{k} \beta_h x_{hij}$$

where the intercept term β_{0j} is assumed to vary randomly and is given by the sum of an average intercept β_0 and group-dependent deviations , u_{0j} that is

$$\beta_{0j} = \beta_0 + u_{0j}$$

As a result,

$$logit(p_{ij}) = \beta_0 + \sum_{h=1}^{k} \beta_h x_{hij} + u_{0j}$$

Solving for p_{ij}

$$p_{ij} = \frac{e^{\beta_0 + \sum_{h=1}^{k} \beta_h x_{hij} + u_{0j}}}{1 + e^{\beta_0 + \sum_{h=1}^{k} \beta_h x_{hij} + u_{0j}}}$$

3.3.2.1.4 The Random Coefficient Logistic Regression Model

This is used to assess whether the slope of any of the explanatory variables has a significant variance component between the groups. Now consider a model with group specific regressions of logit of the success probability,

$$logit(p_{ij}) = log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{0j} + \beta_{1j}x_{1ij} \qquad \text{(a)}$$

The intercepts β_{0j} as well as the regression coefficients, or slopes, β_{1j} are group dependent. These group-dependent coefficients can be split into an average coefficient and the group-dependent deviation

$\beta_{0j} = \beta_0 + u_{0j}$, $\beta_{1j} = \beta_1 + u_{1j}$

Substitution into (a) leads to the model

$$logit(p_{ij}) = log\left(\frac{p_{ij}}{1-p_{ij}}\right) = (\beta_0 + u_{0j}) + (\beta_1 + u_{ij})x_{1ij}$$

$$= (\beta_0 + \beta_1 x_{1ij}) + (u_{0j} + u_{1j}x_{1ij})$$

There are two random group effects, the random intercept u_{0j} and the random slope u_{1j}. It is assumed that the level-two residuals u_{0j} and u_{1j} have means zero given the value of the explanatory variable X. Thus ,β_1 is the average regression coefficient and β_0 is the average intercept. $\beta_0 + \beta_1 x_{1ij}$ is called the fixed part of the model and the second part, $u_{0j} + u_{1j}x_{1ij}$ is called the random part.

These two group effects u_{0j} and u_{1j} will not be independent, but correlated. Further, it is assumed that, for different groups, the pairs of random effects (u_{0j}, u_{1j}) are independent and identically distributed. Thus, the variances and covariance of the level-two random

effects (u_{0j}, u_{1j}) are denoted as follows:

$$var(u_{0j}) = \sigma_{00} = \sigma_0^2$$

$$var(u_{1j}) = \sigma_{11} = \sigma_1^2$$

$$cov(u_{0j}, u_{1j}) = \sigma_{01}$$

The model for a single explanatory variable discussed above can be extended by including more variables that have random effects. Suppose that there are k level-one explanatory variables

$x_1, x_2, \ldots, x_k$, and consider the model where all X-variables have varying slopes and random intercept. That is:

$$logit(p_{ij}) = log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \cdots + \beta_{kj}x_{kij}$$

3.3.2.2 Estimation of the Parameter

The most common methods for estimating multilevel logistic models are based on likelihood. Among the methods, Marginal Quasi Likelihood or MQL [Goldstein (1991), Goldstein and Rasbash (1996)] and Penalized Quasi Likelihood or PQL (Breslow and Clayton, 1993) are the two prevailing approximation procedures. Both MQL and PQL are based on Taylor series expansion to achieve the approximation. Based on the usage of first and second term of Taylor expansion, MQL and PQL are often known as first-order MQL and second-order MQL, first order PQL and second-order MQL respectively. After applying these quasi likelihood methods, the model is then estimated using iterative generalized least squares (IGLS) or reweighted IGLS (RIGLS) [Goldstein (2003)].

Parameter estimation in hierarchical generalized linear models is more complicated than the hierarchical linear models. The most frequently used kind of approximation method used are based on a first-order or second-order Taylor series expansion of the link function. In this study, the multilevel data analysis is supported by the STATA software.

3.3.2.3 Significance testing

As with ordinary least squares regression or logistic regression, we can consider significance tests for individual estimates, such as intercepts, slopes, and their variances, as well as whether the full model accounts for a significant amount of variance in the dependent variable. In between, there is also the possibility of determining whether the subset of predictors contribute significantly.

The fixed effects in multilevel regression are typically tested in a familiar way, by creating a ratio of the intercept or slope estimate to the estimate of the standard error. The usual null hypothesis test is whether the coefficient, either intercept or slope, is significantly different from zero (i.e., is the population value zero or not). This kind of ratio, usually distributed as a z or t, is used in many statistical tests (Newsom, 2011).

$$t = \frac{\hat{\beta}_h}{SE(\hat{\beta}_h)}$$

Where $\hat{\beta}_h$ is either the intercept or slope coefficient and $SE(\hat{\beta}_h)$ is the standard error estimate.

In SPSS (and similarly in SAS, STATA, and MLWIN), fixed effects tests involve the same ratio of the estimate to the standard error estimate, but significance is determined by the normal curve, so it is considered a z-test. The z-test is often referred to as a "Wald" test.

Significance Testing for Random Effects

Random effects tests examine hypotheses about whether the variance of intercept or slopes (or their covariance) is significantly different from zero. The tests of variances and covariance are made using a Wald z-test and chi-square test. The Wald test for variances is simply a ratio of the variance estimate divided by the standard error estimate. Significance tests of variances (but not covariance) using this approach should be interpreted after dividing the p-Value from the output into half (i.e., as a one-tailed test; Snijders & Bosker, 1999).

Likelihood Ratio Tests for Multiple Parameters

Another approach to significance tests involves a comparison of two "nested models." Nested model tests involve comparison of one model to another model that specifies only a subset of the parameters included in the first model (provided the same set of cases are used in both models).

The likelihood ratio test compares the deviance (-2 log likelihood) of two models by subtracting the smaller deviance (model with more parameters) from the larger deviance (model with larger deviance). The difference is a chi-square test with the number of degrees of freedom equal to the number of different parameters in the two models. Any number of parameters can be compared in the two models. In the case where the empty model is compared to a full model, the likelihood ratio test provides information about whether the predictors in the model together account for a significant amount of variance in the dependent variable.

Similarly, the overall model evaluation is also examined using Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC). The smaller the value, the better the model will be fitted.

$$AIC = -2 \times \ln(\text{likelihood}) + 2 \times k$$

$$BIC = -2 \times \ln(\text{likelihood}) + \ln(N) \times k$$

where k is the model degrees of freedom calculated as the rank of variance–covariance matrix of the parameters and N is the number of observations used in estimation or, more precisely, the number of independent terms in the likelihood.

CHAPTER FOUR

STATISTICAL DATA ANALYSIS AND RESULTS

The objective of this chapter is to provide analysis of results about the effect of socioeconomic and demographic characteristics on stunting status of under-five children in rural part of Ethiopia. This chapter is divided into four sections. The first section presents the results of descriptive analysis; the second section, the results of cross tabulation analysis; the third section, analysis of the results obtained from binary logistic regressions; and the last section, analysis of the results obtained from multilevel logistic regression.

4.1. Results of descriptive analysis

The total number of children covered in the study were 7,784. The sample encompasses both male and female children of which, 3,968 children (50.98 %) were males and 3,816 children (49.02 %) were females.

Table 4.1 Results of descriptive analysis of socio-economic and demographic factors

Categories		Height for age status of the child				
		Not stunted	Percentage	Stunted	Percentage	Total
Region	Tigray	472	52.4	429	47.6	901
	Affar	405	51.7	378	48.3	783
	Amhara	511	55.1	417	44.9	928
	Oromia	836	62.5	501	37.5	1337
	Somali	398	72.5	151	27.5	549
	Benishangul-Gumuz	406	55.4	327	44.6	733

	SNNP	761	59.8	511	40.2	1272
	Gambela	438	75.5	142	24.5	580
	Harari	243	71.7	96	28.3	339
	Dire Dawa	207	57.2	155	42.8	362
Mother's education level	No education	3483	58.6	2457	41.4	5940
	Primary education	1137	64.5	627	35.5	1764
	Secondary and above	57	71.2	23	28.7	80
Source of drinking water	Non protected source	2659	59.0	1851	41.0	4510
	Protected source	2018	61.6	1256	38.4	3274
Type of toilet facility	No facilities	4433	59.8	2984	40.2	7417
	Have facilities	244	66.5	123	33.5	367
Number of household members	1-4	1040	59.9	695	40.1	1735
	5-9	3262	59.6	2210	40.4	5472
	10 and above	375	65.0	202	35.0	577
Wealth index	Poor	2594	57.7	1898	42.3	4492
	Medium	916	60.9	587	39.1	1503
	Rich	1167	65.2	622	34.8	1789
Educational level of Husband/partner's	No education	2607	57.9	1895	42.1	4502
	Primary education	1745	61.2	1106	38.8	2851
	Secondary and above	325	75.4	106	24.6	431

Employment status of Mother	No	3397	59.0	2219	41.0	5616
	Yes	1280	62.5	888	37.5	2168
Birth order of the child	1	772	61.7	479	38.3	1251
	2-3	1435	60.6	934	39.4	2369
	4-5	1143	59.3	783	40.7	1926
	6+	1327	59.3	911	40.7	2238
Sex of child	Male	2349	59.2	1619	40.8	3968
	Female	2328	61.0	1488	39.0	3816
Had diarrhea in last two weeks	No	3972	60.7	2577	39.3	6549
	Yes	705	57.1	530	42.9	1235
Had fever in last two weeks	No	3774	60.7	2446	39.3	6220
	Yes	903	57.7	661	42.3	1564
Age of child	0-5 months	766	95.6	35	4.4	801
	6-11 months	628	82.7	131	17.3	759
	12-23 months	779	55.1	636	44.9	1415
	24-35 months	809	53.0	717	47.0	1526
	36-47 months	843	49.7	853	50.3	1696
	48-59 months	852	53.7	735	46.3	1587

The proportion of under-five children stunting varied from one region to another region. The highest proportion (48.3%) of under-five stunted children were recorded in Afar followed by Tigray (47.6%) and Amhara (44.9%). However, the least proportion of under-five stunted children were observed in Gambela (24.5%) followed by Somali (27.5%) and Harari (28.3%).

The prevalence of under-five children stunting were different for different levels of mother's education. About 41.4% of under–five children stunting were recorded for non-educated mothers while 35.5% of the children stunted before the age of 5 were born from mothers with primary education. About 28.7% of stunted children were observed from mothers that have completed secondary school and above. On the other hand the highest proportion of under-five children stunting status (42.1%) were recorded from those who were uneducated fathers compared to 38.8% of under-five children born to primary educated fathers and 24.6% of under-five children stunting status were born from fathers who had completed secondary and higher educational level. Highest prevalence of stunting of under-five children were observed as the educational status decreases, while lowest prevalence of stunting were recorded as the level of education increases in both cases of mother's and husband's or father's educational level.

The experiences of under-five child stunting status was different based on the source of drinking water. Higher proportion of stunting was recorded for children that drunk unprotected source of water and relatively less prevalence of stunting was observed under-five children that drunk protected source of water with a percentage of 41% and 38.4% respectively. The risk prevalence of stunting were high (40.2) for children have no toilet facilities and the proportion of stunted under-five child were (35.5%) recorded as those having toilet facilities.

The proportion (40.4%) of under-five child stunting was observed among households of family size 5-9 and the proportion of under-five child stunting was recorded from households of size 1-4 with percentage of 40.1% and the prevalence of under-five child was observed among households of size 10 and above members which accounted for 35% of under-five children were stunted.

Based on Table 4.1 above, wealth index of parents showed different proportion of under-five child stunting exposure. Children born to poor families had the highest proportion of stunting (42.3% for poor families and 39.1% for families with medium wealth) but children born to rich families had the least proportion of under-five child stunting (34.8% from rich families were recorded).

Table 4.1 also shows that the proportion of children found stunted varies by the employment status of mothers. The higher proportion of the stunted under-five children was from unemployed mothers (41.0 percent). With regards to child age, the highest proportion of stunted under-five children were observed among those whose age group is between 36 and 47 months (50.3 percent) where as the smallest percentage (4.4 percent) of stunted children which was observed among those whose age group is less than or equal to 5 months.

4.2. Results of cross tabulation data analysis

Cross-tabulation analysis, also known as contingency table analysis, is often used to analyze categorical (nominal or ordinal measurement scale) data. A cross-tabulation is a two (or more) dimensional table that records the number (frequency) of respondents that have the specific characteristics described in the cells of the table.

The chi-square (x^2) test is used to assess the relationship between two nominal or ordinal variables. It is a very general statistical test that can be used whenever we wish to evaluate whether frequencies that have been empirically obtained differ significantly from those that would be expected on the basis of chance or theoretical expectations.

Chi-square statistic does not give any information about the strength of the relationship and only conveys the existence or nonexistence of the relationships between the variables investigated.

Based on the result of the cross tabulation analysis, stunting was found to be associated with region (p = 0.000), with mother's educational level (p=0.000), with source of drinking water(p=0.00), with types of toilet facilities (p=0.000), with number of household members (p=0.000) ,wealth index (p=0.000), husband/partner's education (p=0.000), employment status of mother (p=0.000), birth order of the child (p=0.000), sex of child (p=0.001), had diarrhea in last two weeks(p=0.019), had fever in last two week (p=0.000) and with age of child(0.000). (see the Appendix A).

4.3. Logistic regression analysis

4.3.1. Estimation and goodness of fit test

In this section binary logistic regression is applied to assess the relation between stunting status of under-five children, which is a dichotomous response variable, with the explanatory variables.

Multiple logistic regression analysis was applied to assess the effect of each selected variable on stunting, while controlling for other independent variables. The model is fitted to identify the basic demographic, socio-economic, and health and environmental determinants of stunting in under-five children in terms of long-run measures of health outcomes at the rural part of national level. The analysis was done by using STATA with stepwise variable selection method. Table 4.2 shows statistically significant variables and the maximum likelihood estimates of parameters where the significance of each predictor is tested using the Wald test.

Table 4.2 Results of maximum likelihood estimates of parameters in fitting binary logistic regression

Stunting	Coef.	Std. Err.	z	P>z	Odds Ratio	[95% Conf. Interval]	
Region			7.92	0.000*			
Tigray	-1.25166	.042542	-8.42	0.000*	.2860297	.21370	.38283
Afar	-1.038856	.0541231	-6.79	0.000*	.3538593	.26220	.47755
Amhara	-1.213749	.0435781	-8.27	0.000*	.2970813	.22285	.39603
Oromia	-.936	.0546165	-6.72	0.000*	.3921935	.29851	.51527
Benishangul-Gumuz	-1.100901	.0547406	-6.69	0.000*	.3325714	.24086	.45918
SNNP	-.236006	.1160265	-1.61	0.108	.7897759	.59218	1.0533
Gambela	-.4703989	.0849679	-3.46	0.001*	.624753	.47856	.81559
Harari	-1.494736	.0377088	-8.89	0.000*	.2243078	.16134	.31185
Dire Dawa	-.5983334	.0982642	-3.35	0.001*	.5497271	.38725	.78036
Mother's Educational level			7.49	0.000*			
No education	.3868088	.0502884	-7.31	0.000*	1.7327644	1.28615	1.94855
Primary education	.368059	.045572	-8.05	0.000*	1.6343749	1.26502	1.84567
Source of			-6.48	0.000*			

Drinking Water							
Non protected source	.3869562	.1356665	4.20	0.000*	1.4725	1.2292	1.7639
Types of toilet facilities			6.77	0.000*			
No facilities	.366086	.0443392	-8.29	0.000*	1.38057	1.30287	1.57819
Number of household member			-0.85	0.396			
1-4	.2594139	.131899	2.55	0.011*	1.29617	1.0618	1.5822
5-9	.1900621	.1181997	1.94	0.052	1.20932	.99849	1.4646
Wealth index			-4.84	0.000*			
Poor	.4440229	.118029	5.86	0.000*	1.55896	1.3439	1.8083
Medium	.2126466	.1086424	2.42	0.015*	1.23694	1.0413	1.4693
Educational level of husband			-6.71	0.000*			
No education	.7542926	.3046623	5.26	0.000*	2.12610	1.6055	2.8155
Primary education	.6087235	.259754	4.31	0.000*	1.83808	1.3933	2.4246
Employment status of mother			3.87	0.000*			
0	.1793325	.055965	-2.68	0.007*	1.2835279	1.73303	1.95304
BOC			1.60	0.110			
1	-.2766419	.0724671	-2.89	0.004*	.758326	.62880	.91453
2-3	.0782158	.0855192	0.99	0.323	1.081356	.92608	1.2626
4-5	-.2303395	.0668538	-2.74	0.006*	.7942639	.67347	.93672
Sex of child			-3.87	0.000*			
Male	.2140719	.0704092	3.77	0.000*	1.238712	1.1081	1.3846
Had diarrhea in last two weeks			-6.41	0.000*			

No	.5414653	.1375468	6.77	0.000*	1.718523	1.4690	2.0104
Had fever in last two weeks			33.17	0.000*			
No	-2.29179	.0071471	-32.41	0.000*	.1010853	.08800	.11611
Age of child			17.36	0.000*			
0-5 months	-1.841208	.0189013	-15.45	0.000*	.1586257	.12558	.2003
6-11 months	-.4931016	.0620582	-4.85	0.000*	.6107292	.50044	.74531
12-23 months	-.2179007	.0819267	-2.14	0.032*	.8042053	.65864	.98193
24-35 months	-.0865074	.0903305	-0.88	0.380	.9171288	.75612	1.1124
36-47 months	.142605	.1095755	1.50	0.133	1.153274	.95732	1.3893
Constant	1.968399	.2498371	7.88	0.000*		1.4787	2.4580

Note: The symbol “*” indicates that the estimate is significance at =0.05. The reference categories are: "Dire Dawa" for Region, "Secondary and above" for Mother's education, "Improved source" for Source of drinking water, "Have facilities" for Types of toilet facility, "10 and above" for Number of household member, "Rich" for Wealth index, "Secondary and above" for Educational level of husband, "yes" for Employment status of mother, "6+" for Birth order of child, "Female" for Sex of child, "yes" for Had diarrhea in last two weeks, "yes" for Had fever in last two weeks, "48-59 months" for Age of child.

Logistic regression uses maximum likelihood as a method of estimation, and is an iterative procedure. The first iteration (called iteration 0) is the log likelihood of the "null" or "empty" model; that is, a model with no predictors. At the next iteration, the predictor(s) are included in the model. At each iteration, the log likelihood increases because the target is to maximize the log likelihood. When the difference between successive iterations is very small, the model is said to have "converged".

In order to check the goodness-of-fit of an estimated multiple logistic regression model one should assume that the model contains those variables that should be in the model and have been entered in the correct functional form. The likelihood ratio test is used to test the goodness of fit of the model by comparing two nested models:-one with small number of explanatory variables and the other with more explanatory variables. Two models will be compared: one with no variable called empty (intercept only) model and the other with all variables called saturated (full) model.

The deviance (-2LL) is a measure of the difference between a given model and the saturated model, smaller values indicate better fit. we can see that -2LL of the full model is -2(-4,016.4717)= 8,032.9434 and $-2LL_0$ of the intercept only model is -2(-5,236.0351) = 10,472.0702 giving the difference of the deviances is (10,472.0702 - 8,032.9434) =2,439.1268. The model deviance (8,032.9434) is significantly smaller than the null deviance (10,472.0702); hence the set of predictors significantly improved model fit. Since log likelihood ratio test is Chi-square distributed with 31 degrees of freedom, P(x^2 (31) >2439.13)=0.0000, we can reject the null hypothesis of no significant difference between the two models.

The overall goodness of fit of the model was evaluated by Hosmer-Lemeshow goodness of fit test with the null hypothesis:

H_0: the model is a "good enough" fit to the data and we only reject the null hypothesis and decide it is "poor" fit if there are sufficiently strong grounds evidence.

Table 4.3: Hosmer-Lemeshow Goodness of Fit Statistics

Hosmer-Lemeshow goodness of fit test		
Chi-square	DF	Prob>chi2
2.341	8	0.969

The p-value (0.969) is greater than 0.05 in the above table gives us an evidence not to reject H_0. Hence the model is a good fit to the data.

4.3.2 Interpretation of parameter estimates

For a dichotomous variable the parameter of interest is the odds ratio. An estimate of this parameter is obtained from the estimated logistic regression coefficient, regardless of how the variable is coded. This relationship between the logistic regression coefficient and the odds ratio provides the foundation for our interpretation of all logistic regression results. When both the dependent and independent variables are both dichotomous, the odds ratio is the Probability that Y is 1 when X is 1 compared to the probability that Y is 1 when X is 0.

In table 4.2, the column labeled "coefficient" are the values for the logistic regression equation for predicting the dependent variable from the independent variable. These estimates tell the amount of increase in the predicted log odds of the dependent variable (stunting) that would be predicted by a one unit increase in the predictor, holding all other predictors constant.

The Exp (β) (column-6 of table 4.2) gives the odds ratios for each variable. The last category of each explanatory variables were used as reference; subjectively.

The result of binary logistic regression shows that region, mother's education, source of drinking water, facilities of toilet, wealth index, husband/partner's education, employment status of mothers, sex of child, had diarrhea in the two weeks before survey, had fever in the two weeks before survey and age of the child had significant effect on under-five child stunting. In contrast, number of house hold member and birth order of the child had no significant effect on under-five child stunting.

Odds ratio is a measure of association; it is an estimate of the risk of an exposed group relative to control group or unexposed (reference) group. Odds ratio less than 1 indicates negative relationship and odds ratio greater than 1 indicates positive relationship and odds ratio=1 indicates no relationship.

The odds of stunting in Tigray was found to be 0.286 times lower than the odds of stunting in Dire Dawa controlling for the other variables in the model. The odds of stunting decreased by a factor of 0.354 when the person was from Afar compared to a person from Dire Dawa. The odds of stunting in Amhara, Oromia, Somali and Gambela were 0.297, 0.392, 0.333 and 0.224 times lower than the odds of stunting in Dire Dawa respectively; controlling for other variables in the model. The odds of stunting in Ben-Gumuz was 0.789 times lower than the odds of stunting in Dire Dawa. The odds of stunting in Harari and SNNP were 0.549 and 0.625 times lower than the odds of stunting in Dire Dawa respectively; controlling for the other variables in the model.

Children whose mothers had no education were 73.3% (OR=1.733) more likely to be stunted compared to children whose mothers had secondary and above education level controlling for other variables in the model. The odds of stunting of child from mothers who had primary education was 1.634 times higher than the odds of stunting of child from mothers who had secondary and above education controlling for the other variables in the model.

The odds of under-five child stunting that took non protected source of drinking water was 1.472 times higher compared to the odds of stunting of under-five child who took protected source of drinking water controlling for the other variables in the model. On the other hand, children with no toilet facilities (bush/field toilet, hanging toilet, other) were 1.381 more likely to be stunted relative to those who have toilet facilities (pit toilet, flush/pour toilet) controlling for the other variable in the model.

The likelihood of being stunted was 55.89% (OR=1.5589) higher for children from poor families than those from the rich families and children whose families were from medium wealth index were 23.69% (OR=1.2369) more likely to be stunted relative to children whose families were from rich wealth index controlling for other variables in the model. Husband/ partner's educational level were also significantly associated with stunting. For instance, as compared to the reference category (secondary and above educational level), children whose father had no education and primary level of education were 2.126 and 1.838 more likely to be stunted respectively; controlling for other variables in the model.

Children whose mother is unemployed were 28.3% more likely to be stunted compared to employed mother controlling for other variables in the model. On the other hand, the odds of stunting for female under-five child was 1.239 times higher than the odds of the stunting for male controlling for other variable in the model.

The model shows that the odds of stunting for children who had no diarrhea two weeks before the survey were 1.7185 times the odds of stunting for children who had diarrhea two weeks before the survey. In other words, children who had no diarrhea are 71.85% more likely to be stunted than children who had diarrhea two weeks before the survey controlling for other variables in the model. The odds of stunting for children who had no fever two weeks before the survey date were 0.101 times the odds of stunting for children who had fever two weeks before the survey. In other words, children who had no fever are 10.1% less likely to be stunted than children who had fever two weeks before the survey controlling for other variables in the model.

The exposure of stunting on under-five children with age between 0-5 months was 0.1586 times lower than the risk of stunting on under-five children with age between 48-59 months and the odds of stunting for under-five children with age between 6-11 months, 12-23 months, and 24-35 months were 0.611, 0.804 and 0.917 times lower than the odds of stunting for the reference category (age between 48-59 months) respectively; controlling for other variables in the model. The risk of stunting on children with age between 36-47 months was 15.3% more than the risk of stunting on children with age between 48-59 months controlling for other variables in the model.

4.3.3. Model diagnostics

After fitting a regression model it is important to determine whether all the necessary model assumptions are valid before performing inference. If there are any violations, subsequent inferential procedures may be invalid resulting in faulty conclusions. Therefore, it is crucial to perform appropriate model diagnostics. The results of the model diagnostic are shown in the Appendix B.

All the values of the model diagnostic measures are less than unity. A DFBETAs less than unity indicates that no one value has an effect on the estimate of a regression coefficient of a particular predictor variable. Cook's distance less than unity implies that no one observation has an effect on a group of regression coefficients β. A value of the leverage statistic less than unity confirms that no observation is far apart from the others in terms of the levels of the independent variables. Based on the above goodness of fit tests and diagnostic checking results, we can decide that our model is adequate.

4.4. Multilevel logistic regression analysis

The data used in this study have a hierarchical structure. Units at one level are nested within units at the next higher level. Here, the lower level (level-1) units are the individual children, and the higher level (level-2) units are the regions that constitute the groups into which the children are clustered or nested. The nesting structure is children within regions that resulted in a set of 10 regions with a total of 7,784 children. The data used in this study consist of variables describing individuals as well as variables describing regions. Therefore, the statistical model used has to describe the data at both levels, in order to find the effect on stunting of both the individual children and the regions.

As mentioned above, multilevel models were developed to analyze hierarchically structured data. These models contain variables measured at different levels of the hierarchy. Unobserved heterogeneity is modeled by introducing random effects. Random intercepts are used to model unobserved heterogeneity in the overall response; random coefficients model unobserved heterogeneity in the effects of explanatory variables on the response variable. As one of the aims of this study was to model the heterogeneity between regions, a random intercept model was used. This allows that the overall probability vary across the regions. The chi-square test was applied to assess heterogeneity between regions. The test yields x^2 = 178.0768 with df 9 (P=0.000). Thus, there is evidence of heterogeneity among the regions.

Children in this study were selected from different regions of Ethiopia. Thus there are two kinds of random variability in the data, variability between different children in a single region, and variability between different regions. The advantages of using a multilevel model include the ability to fully explore the variability at all levels of the data hierarchy, and estimation of correct standard errors in the presence of clustered data. The model takes into account the correlation structure of the data, enabling correct inferences to be made.

4.4.1 Analysis of the Empty Model with Random Intercept

This is the type of model that incorporates only the grand mean and random intercept (regional effect) without covariate. The model is given as:

$$logit(\pi_j) = f(p) = \beta_0 + U_{0j}$$

The intercept β_0 also known as the grand mean is shared by all regions while the random effect U_{0j} also known as level two residual is specific to region j. It shows how the mean of stunting in a particular region deviates from the grand mean. σ_u^2 is the between regions variance. The random effect is not directly estimated but is summarized in terms of their estimated variances.

Table 4.4. The empty model with random intercept

STUNT	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
_cons	-.4758037	.115806	-4.11	0.000	-.7027792	-.2488282

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
RGN: Identity			LB	UB
var(_cons)	.1271605	.0606404	.049938	.3237976
Chi2(01) = 141.37 p-value = 0.0000				

Model	Obs	LL(model)	df	AIC	BIC
Ch	7784	-5165.351	2	10334.7	10348.62

Table 4.4 shows the output of the estimates of fixed effects and random effects. From the table we can see that the estimate of the fixed part of the model is -0.4758 with z-value of -4.11 and p-value of 0.000 which implies that the average log odds of stunting is significantly different from zero. The fixed part of the model is interpreted as the grand mean of log odds of stunting with the odds of exp (-0.4758) = 0.621 and the average probability of under-five child stunting $\frac{\exp(-0.4758)}{1+\exp(-0.4758)} = 0.383$ which means the chance of under-five child stunting is 38.3% on average. The table also contains the variance estimate of random effects at regional level, σ_u^2= 0.1272 with confidence interval of (.0499, .324) which implies that the between region variance of under-five child stunting is 0 .1272. At the bottom of the second table the result of the hypothesis H_0: σ_u^2 = 0 is provided showing that there is no cross-regional variation in under-five children stunting. For this hypothesis, we see that the value of the test statistic is 141.37 with p = 0.000*. Therefore, the null hypothesis is rejected and there is evidence of heterogeneity or cross regional variation in under- five children stunting.

We can now write the model for the j^{th} region as logit(π_j) = -0.4758 + u_{0j}. From the model we can say that the average probability of under-five child stunting in the absence of covariates in region j is less than the average when u_{0j} is negative while it is higher than the average when u_{0j} is positive. The variance of the random factor in the empty model is significant which indicates that there are significant regional differences in the height-for-age status of under-five children. Based on the significant LR the chi-square value for the no intercept model implies that an empty model for stunting with random effect is better than an empty model for stunting without random effect.

4.4.2 Random Intercept Model with Fixed Coefficient Model

In random intercept and fixed slope model, covariates are included but none of them is allowed to have a cluster-specific effect upon the response, i.e. each covariate's effect is assumed to be the same in the clusters. The probability of stunting is allowed to vary across regions while level-one covariates including the fixed intercept are fixed or constant across regions. The results of the random intercept and fixed slope model are presented in Table 4.5.

Table 4.5. The random intercept with fixed slope model

Wald chi2(22) = 1564.28 Prob > chi2 = 0.0000

Stunting	Coefficients	Std. Err.	Z	P>z	Odds Ratio	[95% Conf. Interval]	
Mother's educational level							
No education	.3788	.05055	-7.28	0.000	1.75743	1.288648	1.48911
Primary education	.3406	.04574	-8.03	0.000	1.54573	1.26672	1.80622
Source of drinking water							
Non protected source of water	.39002	.13591	4.24	0.000	1.4770	1.23327	1.7689
Types of toilet facilities							
No facilities	.46525	.04429	-8.30	0.000	1.38088	1.30324	1.7840
Number of household member							
1-4	.26308	.13223	2.59	0.010	1.3009	1.0659	1.5877
5-9	.19224	.11833	1.97	0.049	1.2119	1.0008	1.4675
Wealth index							

Poor	.43724	.11691	5.79	0.000	1.5484	1.3354	1.7953
Medium	.21258	.10849	2.43	0.015	1.237	1.0418	1.4692
Educational level of husband							
No education	.7638	.30688	5.34	0.000	2.1464	1.6219	2.8406
Primary education	.61964	.26221	4.39	0.000	1.8582	1.40923	2.4502
Employment status of mother							
No	-.18269	.05570	-2.73	0.006	.83307	.73073	.94974
Birth order of child							
1	-.27660	.07239	-2.90	0.004	.75835	.62894	.91439
2-3	.07809	.08543	0.99	0.323	1.0812	.92608	1.2623
4-5	-.23022	.06680	-2.74	0.006	.79435	.67365	.93669
Sex of child							
Male	.21319	.07029	3.75	0.000	1.2376	1.1072	1.3833
Had diarrhea in last two week							
No	.54015	.13712	6.76	0.000	1.7162	1.4675	2.0072
Had fever in last two week							
No	-2.2779	.00722	-32.30	0.000	.10249	.08926	.11768
Age of child							
0-5 months	-1.839	.01891	-15.45	0.000	.15887	.12581	.20064
6-11 months	-.4914	.06210	-4.84	0.000	.61172	.50134	.74640
12-23 months	-.2169	.08193	-2.13	0.033	.80500	.65941	.98273
24-35 months	-.0866	.09023	-0.88	0.378	.91697	.75612	1.1120
36-47 months	.1488	.10969	1.52	0.130	1.1547	.95870	1.391
_cons	1.1092	.25438	4.36	0.000		.61065	1.6078

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]

Region: Identity				
var(constant)	.196983	.0935714	.0776406	.4997684
Chi2(01) = 155.62 p-value = 0.000				

The reference categories are: "Dire Dawa" for Region, "Secondary and above" for Mother's education, "Improved source" for Source of drinking water, "Have facilities" for Types of toilet facility, "10 and above" for Number of household member, "Rich" for Wealth index, "Secondary and above" for Educational level of husband, "yes" for Employment status of mother, "6+" for Birth order of child, "Female" for Sex of child, "yes" for Had diarrhea in last two weeks, "yes" for Had fever in last two weeks, "48-59 months" for Age of child.

The Wald test of overall goodness of fit with Wald chi2 (22) =1564.28 and p =0.000, where 22 is the degrees of freedom, indicates that all explanatory variables jointly are significant. The variance component representing variation between regions has increased from 0.127 in the empty model with random intercept to 0.197 in the random intercept with fixed slopes multilevel logistic regression model. This indicates that there is significant variation between regions in the stunting status of children.

The values of chi2 (1) =155.62 and p=0.000 in the last column of the table leads to the rejection of the null hypothesis that the random effect is zero as in the assumption of ordinary logistic regression. From this we can conclude that the random effect at regional level is significantly different from zero.

Table 4.6. Model comparison criteria between the empty and the random intercept models

	Empty model with random intercept	Random intercept with fixed effects model
Log likelihood (LL)	-5165.351	-4037.9364
Deviance = -2LL	10330.702	8075.8728
AIC values	10334.7	8123.873
BIC values	10348.62	8290.909

Likelihood-ratio test	LR Chi2(22) = 1564.28	Prob>chi2 = 0.000

The empty model with random intercept and the random intercept model with fixed coefficients are compared in Table 4.6. The deviance of the empty model with random intercept (10330.702) is greater than the deviance of the random intercept model with fixed coefficients (8075.8728). The smaller the deviance the better the model; hence random intercept model with fixed coefficients is the better fit to the data.

Moreover, in table 4.6, both the AIC and BIC of the empty model with random intercept are greater than that of the random intercept model with fixed coefficients. The smaller the AIC and BIC of a model the better the fit is. In all criteria the random intercept model with fixed coefficient the better model.

4.4.3 Random Intercept Model with Random Coefficients

Multilevel logistic regression can allow the coefficient of level-one covariates to vary across regions instead of keeping them fixed across regions. In random intercept model with random coefficients the children's level covariates are allowed to vary randomly across regions.

This model contains fixed effects and random effects. The fixed effects are analogous to the standard logistic regression coefficients and are estimated directly. The random effects are not directly estimated but are summarized in terms of their estimated variances and covariance. The random effects can take random intercepts (regional effects) and random coefficients (level-one covariates effect). In this section we investigate whether level-one covariates have random or fixed effects across regions. All variables included in the random intercept model are included in random coefficient model. Estimates of this model showed that the variance of random slopes of all included variables are zero except for educational level of mother and age of child. This indicates that only the effects of these variables varied across regions whereas the effect of other covariates for stunting remained fixed across regions. The results of the random slope estimates are given in Table 4.7 below.

Table 4.7. The random coefficients model

Wald chi2(22) = 1346.16 Prob > chi2 = 0.0000

Stunting	Coefficients	Std. Err.	Z	P>z	Odds Ratio	[95% Conf.	Interval]
Mother's educational level							
No education	.3804	.12038	-2.80	0.005*	1.727	1.0806	1.9418
Primary education	.2125	.084825	-4.22	0.000*	1.586	1.1599	1.7511
Source of drinking water							
Non improved source	.4567	.132585	3.84	0.000*	1.428	1.1911	1.7137
Types of toilet facilities							
No facilities	.40383	.05603	-6.42	0.000*	1.447	1.3502	1.5720
Number of household members							
1-4	.2669	.13464	2.59	0.010*	1.3059	1.0670	1.5984
5-9	.12516	.11253	1.26	0.207	1.1333	.93291	1.3768
Wealth index							
Poor	.3846	.11420	4.95	0.000*	1.4690	1.2614	1.7108
Medium	.1881	.10971	2.07	0.038*	1.2070	1.0100	1.4424
Educational level of husband							
No education	.82613	.33391	5.65	0.000*	2.2844	1.7154	3.0423
Primary education	.65681	.2773	4.57	0.000*	1.9286	1.4550	2.5564
Employment status of mother							

No	-.1172	.0613	-1.70	0.089	.8893	.77694	1.018
Birth order of child							
1	-.31606	.0712	-3.24	0.001*	.7290	.60199	.8828
2-3	.05385	.08578	0.66	0.508	1.0553	.8999	1.237
4-5	-.21783	.0695	-2.52	0.012*	.80426	.6788	.95278
Sex of child							
Male	.23415	.07358	4.02	0.000*	1.2638	1.1275	1.4166
Had diarrhea in last two weeks							
No	.53855	.1386	6.66	0.000*	1.7135	1.4622	2.0079
Had fever in last two weeks							
No	-2.2160	.00769	-31.41	0.000*	.10904	.09496	.12521
Age of child							
0-5 months	-1.9310	.03819	-7.33	0.000*	.14499	.08652	.24298
6-11 months	-.53327	.12579	-2.49	0.013*	.58668	.3853	.89313
12-23 months	-.23294	.13838	-2.33	0.018	.79219	.5625313	1.1156
24-35 months	-.08307	.12623	-1.61	0.054	.92028	.70333	1.2041
36-47 months	.15090	.12496	2.40	0.016	1.1628	.94202	1.4355
constant	1.37537	.6254363	2.20	0.028		.14953	2.6012

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
RGN: Unstructured				
var(Mother's education)	.6307895	.3016201	.2470993	1.610265
var(Age of child)	.0207942	.0111963	.0072382	.0597388
var(constant)	.2792657	.1506924	.0969856	.8041329
cov(Mother's education, Age of child)	.100369	.0531764	-.0038548	.2045929
cov(Mother's education, constant)	-.361755	.1948406	-.743636	.020125
cov(Age of child, constant)	-.053067	.0356309	-.1229024	.0167681
LR Chi2(6) = 415.90 p-value =0.0000				

Note: ref indicates reference and '*' indicates significance at 5% significance level. The reference categories are: "Dire Dawa" for Region, "Secondary and above" for Mother's education, "Improved source" for Source of drinking water, "Have facilities" for Types of toilet facility, "10 and above" for Number of household member, "Rich" for Wealth index, "Secondary and above" for Educational level of husband, "yes" for Employment status of mother, "6+" for Birth order of child, "Female" for Sex of child, "yes" for Had diarrhea in last two weeks, "yes" for Had fever in last two weeks, "48-59 months" for Age of child.

Wald chi2 (22) = 1346.16 with p-value=0.000 indicates that at least one population parameter is significantly different from zero. Moreover, the likelihood ratio (LR) test shown at the foot of the table provides evidence that the logistic regression model with no intercept is rejected by the data. This means that the variance-covariance of random effects of the population is significantly different from zero.

The estimate of the fixed intercept is 1.375 and the log-odds of the probability of stunting when all level one covariates are zero in region j is given by 1.375 + $\hat{u}_{j,}$ where $\hat{u}_j$ is a random intercept with variance of 0.279 (indicated in the table as var(cons)) which is the between-regions variance and standard error 0.151. In the absence of level-one covariates, the status of each region on stunting as compared to the average child stunting measured with log odds depends on the sign of the random intercept, $\hat{u}_j$. When $\hat{u}_j$ is positive the log odds of under-five child stunting is higher than the average and when $\hat{u}_j$ is negative the log odds of under-five child stunting is less than the average. The individual region slopes of mother's education level (ME) and age of child (AC) vary with variance 0.631 and 0.021, respectively.

The correlation matrix contains the estimated correlation between random intercepts and the random slopes. Positive correlation between intercepts and slopes implies that regions with higher intercepts tend to have on average higher slopes. The negative sign for the correlation between intercepts and slopes implies that regions with higher intercepts tend to have on average lower slopes on the corresponding predictors (see Table 4.8).

Table 4.8 Random-effects correlation matrix for level REGION

	Mother's education	Age of child	Constant
Mother's education	1		
Age of child	0.876	1	
Constant	-0.862	-0.696	1

In order to see whether the inclusion of level one covariates; mother's educational level and age of child varying across regions significantly improved the random intercept only model, log likelihood ratio test, AIC and BIC values have been used. The likelihood ratio test results are LR chi2 (22) = 2515.104 (which is the difference between the deviance of random intercept only model and random intercept with random slopes model ; 10330.702- 7815.598) and p= 0.000 implying that there is a significant difference between the two nested models. The AIC(7873.597) and BIC(8075.432) values for the random intercept with random slope model are less than the corresponding AIC(10334.7) and BIC(10348.62) values for the random intercept only model respectively; indicating that the random intercept with random coefficient model is better than random intercept only model.

As discussed earlier, the random intercept model with fixed coefficients and the random intercept model with random coefficient are found to be better than that of the empty model. Below we compare the random intercept model with fixed coefficients and the random intercept model with random coefficient.

Table 4.9 results of model selection criteria and log likelihood ratio test

Model Selection Criteria	Random intercept with fixed effect model	Random intercept with Random coefficient model
Log likelihood	-4037.9364	-3907.799
Deviance	8075.873	7815.598
AIC	8123.873	7873.597
BIC	8290.909	8075.432

Both the AIC and BIC values of the random intercept with fixed coefficient model are greater than the random intercept with random coefficient model. The deviance of random intercept with random coefficient model less than that of random intercept with fixed effect model. In all the three criteria as shown in Table 4.9 random intercept with random coefficient model found to be better model than random intercept with fixed effect model . Hence, the random intercept with random coefficient model is the best model.

The parameters of observed variables can be interpreted much the same way as those from the standard logit model. Thus, everything else being equal except slight difference on random effect in the model, children whose mother have no education were 72.75% more likely to be stunted (OR=1.7275) compared to children whose mother have secondary and above educational level controlling for other variables in the model and random effect at level two while children whose mother had primary education were 58.6% (OR=1.586) more likely to be stunted compared to children whose mother have secondary and above education controlling for other variables in the model and random effect at level two.

Children who live in poor households were about 46.9% more likely to be stunted than that of children who live in rich households controlling for other variables in the model and random effect at level two. The odds of being stunted for a child who drunk unimproved source of water were 42.87% more likely to be stunted than those children who drunk improved source of water controlling for other variables in the model and random effect at level two.

The odds of stunting of children who had toilet facilities were 1.4476 (OR=1.4476) times higher than the odds of stunting of children with no toilet facilities controlling other variables in the model and random effects at level two. On the other hand, The odds of stunting for children who had no fever two weeks before the survey date were 0.109 times the odds of stunting for children who had fever two weeks before the survey.

The odds of stunting status of children whose mother are unemployed were lower by a factor of 0.889 compared to children whose mother are employed controlling for other variables in the model and random effects at level two.

CHAPTER FIVE

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

In this chapter we discuss the main findings in this study and also present the conclusions drawn and recommendations made.

5.1 Discussion

The outcome of this study shows that stunting of under-five children in rural parts of Ethiopia is very high. What is probably more important is to come up with some suggestions that will hopefully help alleviate the problem of stunting in rural part of Ethiopia.

This study intended to identify significant socio-economic and demographic determinants of under-five child stunting based on the 2011 EDHS data. The descriptive analysis results indicated that under-five child stunting varied among the regions. The highest under-five child stunting was observed in Afar followed by Tigray and the least under five child stunting was observed in Gambela followed by Somali. This disparity could be due to unequal distribution of infrastructures and different socio-economic and demographic characteristics exercised in each region. The higher under-five child stunting status was observed from children that drink not protected sources of drinking water compared to stunting of under-five child who drink protected sources of drinking. On the other hand, children who used toilet with no facilities were more likely to be stunted relative to the reference group (have toilet facilities). The exposure of stunting was significantly higher for children whose mothers were unemployed than children whose mother were employed.

The risk of stunting was significantly higher for children whose mothers had no education and primary education level than children whose mothers had secondary and above level of education. This finding seemed to be consistent with other studies (Oyekale and Oyekale, 2000; Smith and Haddad, 2000). They indicated that education improves the ability of mothers to implement simple health knowledge and facilitates their capacity to manipulate their environment including health care facilities, interact more effectively with health professionals, comply with treatment recommendations, and keep their environment clean. Furthermore, educated women have greater control over health choices for their children. The result of the current study is also congruous with a study done in Bangladesh. The analysis of that study indicated that the educational level of caregiver was positively related to the better nutritional status of children. This is likely to be attributed to better education because educated mothers are more conscious about their children's health; they tend to look after their children in a better way (Rayhan and Hayat, 2006).

Rahman et al (2009) made a study based on the experience of 63 developing countries over 25-years period to identify the determinants of child malnutrition for each developing region. Six factors were explored; one of the important factors was women's education. They depicted that improvements in female secondary school enrollment rates were estimated to be responsible for 43% of the total 15.5% reduction in the child malnutrition rate of developing countries during the period 1970-95 (Rahman et al, 2009). In addition, the percentage of stunting amongst fathers with low educational levels was the highest, whereas fathers who are illiterate have stunted children in a percentage much higher than those of secondary and above educational level. The percentage of stunting in fewer than three children is relatively identical regarding fathers with primary educational level. This fact may be related to inability of illiterate or of low educational level of fathers to meet his children nutritional needs, because of the low income, and increase in the percentage of unemployment fathers.

Our study also revealed that under-five children from poor households are at a higher risk of stunting than children from rich households. This finding is consistent with other studies (Smith et al., 2005; Woldemariam and Timotewos, 2002). They indicated that rich households have better access to food and higher cash incomes than poor households, allowing them a quality diet, better access to medical care and more money to spend on essential nonfood items such as schooling, clothing and hygiene products.

The result of the present study indicates that age of child is also one of the important determinant factors of stunting of under-five children in rural part of Ethiopia. The prevalence of stunting was highest in children aged between 36 and 47 months than the other age groups. The findings of this study also show that children who had no fever two weeks before date of survey are significantly vulnerable to stunting than those who had not. This finding is consistent with other studies (Sommerfelt et. al., 1994; WHO, 1986). Similarly, the study also showed that children who had no diarrhea two weeks before the survey are significantly vulnerable to stunting than those who had not.

This study examined individual and community-level (regional) factors as significant determinants of childhood stunting in under-five children. It confirms the importance of regional variations with respect to childhood stunting.

Using multilevel logistic regression method of analysis, this study examined variations in childhood stunting among geographical regions. The model suggests that the status of stunting of children differs among regions, although the variations among different communities/regions with respect to the odds of having childhood stunting were found to originate mainly from variations in individual-level factors. These findings are consistent with most studies that have tried to differentiate contextual effects from compositional effects (Frohlich et al. 2002; Subrama-nian et al. 2003) and support a major role for community-level phenomenon as a strong influence on childhood stunting.

5.2 Conclusions

The study identified the following socio-economic, demographic and health and environmental variables as important determinants of stunting of under-five children in rural part of Ethiopia: Age of child, mother's educational level, wealth index, employment status of mother, source of drinking water, types of toilet facilities, had diarrhea in the two weeks before survey, had fever in the two weeks before survey, educational level of partner and geographical region. Childhood malnutrition remains a major public health problem in rural part of Ethiopia.

Although there is regional disparity in children's health/nutritional status, it is observed that children living in rural parts of the country were at high risk of being stunted. The result also suggested that children from poor families were more likely to be stunted than children from rich families in rural Ethiopia.

5.3 Recommendations

Based on our findings, we recommend the following:

- Create awareness about child feeding practice in scientific way.
- Adult literacy programs with special focus on child health and nutrition should be organized particularly for women in communities with a high illiteracy rate as a short term solution aimed at increasing low literacy level in rural Ethiopia.
- Any intervention by governmental and non-governmental organizations that aims at improving under-five children's nutritional status should consider regions with high rates of childhood stunting so as to avert under-coverage of the regions that deserve it.
- Design and implement primary health care and nutrition programs which would fit the features of each region to safeguard children from nutritional deficiency.

REERENCE

Abay Asfaw (1995). How Poverty Affects the Health Status and the Health Care Demand Behavior of Households? The Case of Rural Ethiopia. Agricultural Economics**,** Vol. 30 (3): 215-228.

Abey Koon (1995). Sex preference in South Asia: Srilanka and outliers*.* Asia-Pacific Population Journal 32(1):69-91.

Administration Committee on Coordination–Sub-Committee on Nutrition (ACC/SCN) (2000). Fourth Report on the World Nutrition Situation: ACC/SCN in Collaboration with International Food Policy Research Institution, Geneva.

Agresti, A. (2002). An Introduction to Categorical Data Analysis. 2nd Edition John Wiley and Sons Inc., New York

Ajieroh, V. (2010). *A Quantitative Analysis of Determinants of Child and Maternal Malnutrition in Nigeria*. IFPRI Nigeria Strategy Support Program Brief No. 11, 2010.

Alderman, H., Behrman, J., and Hoddinott, J. (2004). Nutrition, malnutrition and economic growth in Health and economic growth: findings and policy implications. López-Casasnovas,G., Rivera, B. andCurrais, L., Cambridge,MA: MIT Press

Alderman, H., Hoddinott, J. and Kinsey B. (2002). Long term consequences of earlychildhood malnutrition. Mimeo, Department of Economics, Dalhousie University, Halifax, Canada.

Alemu Mekonnen, Bekele Tefera, Tassew Woldehanna, Jones, N., Seager, J., Tekie Alemu, Getachew Asgedom (2005b). Child Nutritional Status in Poor Ethiopian Households: The Role of Gender, Assets and Location. Working Paper No. 26, Young Lives, Save the Children UK.

Alemu Mekonnen, Jones, N., Bekele Tefera (2005a). Tackling Child Malnutrition in Ethiopia: Do Sustainable Development Poverty Reduction Programme's Underlying Policy Assumptions Reflect Local Realities? Working Paper No. 19, Young Lives, Save the Children UK.

Basu, K. and Foster, J. (1998). On measuring literacy, Economic Journal, 108:1733-49 Bender, D.A.(1997). Introduction to Nutrition and Metabolism (2nd Ed.). London: Taylor and Francis.

Bilisuma Bushie (2004). Determinants of Child Stunting in Urban Ethiopia. Unpublished Msc Thesis. AAU, June.

Blossner, M. and de Onis, M. (2005). Malnutrition: Quantifying the Health Impact at National and Local Levels. Environmental Burden of Diseases Series, No. 12.

Breslow, N. E. and D. G. Clayton (1993). Approximate inference in generalized linear mixed models. J. Am. Statist. Assoc.

Butterworth, C.E., "The Skeleton in the Hospital Closet," *Nutrition Today* 9 (1974):4–8. Cassens, Digna, "Enhancing Taste, Texture, Appearance, and Presentation of Pureed Food: Improved Resident Quality of Life and Weight Status," *Nutrition Reviews* 54 (1996):552.

Christiaensen, L. and Alderman, H. (2004). Child malnutrition in Ethiopia: can maternal knowledge augment the role of income: Economic Development and Cultural Change, **52(2)**:287-312.

Collier, P. Dercon, S. and Mackinnon, J. (2002). "Density Versus Quality in Health Care Provision: Using Household Data to Make Budgetary Choices in Ethiopia". The World Bank Economic Review, Vol. 16, No. 3, 425-448.

Cook, R. and Weisberg, S. (1982). Residuals and Influence in Regression. Chapman and Hall, New York.

Copas, J.B. (1988). Binary Regression Models for Contaminated Data. Journal of Royal Statistical Association, B 50 (2).

De Onis M, Monteiro C, Akre, J. and G., Clugston (2003).The worldwide magnitude of protein-energy malnutrition: an overview from the WHO Global Database on Child Growth.

EDHS (2011).Ethiopian Demographic and Health Survey Central Statistics Agency, Addis Ababa, Ethiopia.

FAO. (2008). *The State of Food Insecurity in the World 2008*. Food and Agriculture Organization, Rome.

Federal Democratic Republic of Ethiopia (FDRE), (2005) "National Nutrition Strategy For Ethiopia." Unpublished. Addis Ababa.

Frohlich K.L., Potvin L., Gauvin L. & Chabot P. (2002) Youth smoking initiation: disentangling context from composition. Health & Place 8, 155–166. Available from.

Genebo, T., Girma, W., Hadir, J. and Demmissie, T. (1999).The Association of Children's Nutritional Status to Maternal Education in Ziggbaboto, Guragie Zone South Ethiopia. Ethiopian Journal of Health Development 13(1):55-61.

Genebo, T and Girma, W. (2002). Determinants of nutritional status of women and children in Ethiopia. Calverton, Maryland: ORC Macro.

Getaneh, T., A. Assefa, and, Z. Taddesse. 1998. Protein energy malnutrition in urban children: Prevalence and determinants. *Eth. Med. J.* 36(3).

Gibson,J. (2001). Literacy and intra-household externalities, World Development, **29**: 155-166.

Glewwe, P. (1999). Why does mother's schooling raise child health in developing countries? Evidence from Morocco, Journal of Human Resources, **34(1)**: 124-159

Goldstein, H. (2003). Multilevel Statistical Models. (3rd Ed.). London: Arnold; New York: Oxford University Press Inc.

Goldstein, H. and J. Rasbash., (1996). Improved approximations for multilevel models with binary responses. J. Roy. Statist. Soc. A.

Goldstein, H., (1991). Nonlinear multilevel models with an application to discrete response data. Biometrika, **78**,45–51.

GugsaYimer (2000). Malnutrition among Children in Southern Ethiopia: Levels and Risk Factors. Ethiopian Journal of Health Development, 14(3):283-292.

Hosmer, W.D. and S. Lemeshow (2000). Applied Logistic Regression. 2nd Ed., John Wiley and Sons, New York.

Jeyaseelan, L. 1997. Risk factors for malnutrition in south India children. *Journal of Biosocial Science* 1: 93-100.

Kikafunda Joyce K.,Walker Ann F., Collett David and Tumwine James K. (1998). "Risk factors for early childhood malnutrition in Uganda". *Official journal of the American academy of pediatrics*.Vol.102; P.1_ 8.

Lang Susan S. (May 29, 1998). "Stunted growth affects almost 40 percent of the developing world's infants". Cornell news.P-30.
http://www.news.cornell.edu/releases/May98/stunting.ssl.html

Lewit Eugene M., Kerrebrock Nancy. (1997). "Population based growth stunting". The future of children.Vol.7, no.2; P.149_156.

Moen, A.A. (1993). The impact of child and maternal survival programmes on socioeconomic development, Center for Economic Research on Africa, School of Business, Montclair State University and Upper Montclair, New Jersey.

Nuha Mamoun, Susan Homedia, Mustafa Mabyou and Hussan M. Ahmed Muntasir T. Salah and Ishag Adam Prevalence, Types and Risk Factors for Malnutrition in Displaced Sudanese Children. The Academy of Medical Sciences and Technology, Khartoum, Sudan College of Medicine, International University of Africa Ribat University, Sudan Faculty of Medicine, University of Khartoum, Sudan.

Newsom (2011). Significance Testing in Multilevel Regression.

Oyekale, A. S.and Oyekale, T. O. (2000). Do Mother's Education Levels Matter in Child Malnutrition and Health Outcomes in Gambia and Niger, Department of Agricultural Economic, University of Ibadan, Nigeria.

Olivieri, F., Semproli, S., Pettener, D., & Toselli, S. (2008). "Growth and malnutrition of rural Zimbabwean children (6-17 years of age)".*American Journal of Physical Anthropology*.Vol. 136, Issue 2, P. 214 ñ 222.

Padmadas Sabu S, Hutter Inge and Willekens Frans. "Weaning initiation patterns and subsequent linear growth progression among children aged 2 and 4 years in India". (2002). *International Journal of Epidemiology.*Vol.31. P.855_863.

Rahman Mosiur, Mostofa Golam and Nasrin Sarker Obaida. (2009). "Nutritional status among children aged 24-59 months in rural Bangladesh". *The Internet Journal of Biological Anthropology.* Vol. 3, No.1.

Rayhan Israt and Hayat Khan Sekander. (2006). "Factors Causing Malnutrition among under Five Children in Bangladesh". *Pakistan Journal of Nutrition.* Volume 5, P. 558-562.

Ricci Judith A and Becker Stan. (1996). "Risk factors for wasting and stunting among children in Metro Cebu, Philippines". *American journal of clinical nutrition.* Vol. 63. P. 966-975.

Rice Amy L., Sacco Lisa, Hyder Adnan, and Black Robert E. (2000). "Malnutrition as an underlying cause of childhood deaths associated with infectious diseases in developing countries". Bulletin of the World Health Organization. Vol.78, No.10: P. 1207-1221.

Rolfes Sharon Rady, Pinna Kathryn, and Whitney Ellie. (2004). "Understanding Normal and Clinical Nutrition". National academic of Sciences. 7th edition. P. 522_534.

Shaikh S, Mahalanabis D, Chatterjee S, Kurpad A V and Khaled M A (2003). "Lean body mass in preschool aged urban children in India: gender difference". European *Journal of Clinical Nutrition.* Vol. 57, P. 389ñ393.

Sentayehu Gebre Giorgis (1994). Determinants of the Health and Nutritional Status of Children in Rural Ethiopia: The Case of Sidamo. Unpublished Msc Thesis. AAU, July.

Shrimpton Roger and Kachondham Yongyout. (2003). "Analyzing the Causes of Child Stunting". P. 4-8.

Silva, P. (2005). Environmental Factors and Children's Malnutrition in Ethiopia. World Bank Policy Research Working Paper, 3489.

Smith, L.C. and Haddad, L. (2000). Explaining Child Malnutrition in Developing Countries: A Cross-Country Analysis." IFPRI Research Report 111. Washington D.C. International Policy Research Institute.

Smith, L. C., Ruel, M. T. and Ndiaye, A. (2005**).** Why is Child Malnutrition Lower in Urban than in Rural Areas? Evidence from 36 Developing Countries' World Developmen*t*, Vol. 33, No. 8, 1285-1305.

Snijders, T. A. B. and Bosker, R. J. (1999). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. SAGE Publications, London.

Sommerfelt, A. Elizabeth, and S. Kathryn (1994**).** Children's nutritional status. DHS Comparative Studies No. 12.Calverton, Maryland, USA: Macro International Inc.

Stalin P., Joy Bazroy, Dinesh Dimri, Zile Singh and Senthilvel V. (2013). Prevalence of Stunting and its Risk Factors among Under Five Children in a Rural Area of Kancheepuram District in Tamil Nadu, India (Issue 6 (Jan.- Feb.), **3**, 71-74).

Subramanian S.V., Lochner K.A. and Kawachi I. (2003) Neighborhood differences in social capital: a compositional artifact or a contextual construct. Health & Place 9, 33–44. Available from.

United Nations Children's Fund (UNICEF) (2007). The State of the World's Children 2007: Women and Children. UNICEF, New York.

Van de Poel, E., Hosseinpoor, A., Speybroeck, N., Van Ourti, T., Vega, J., 2008. Socioeconomic inequality in under-five children malnutrition in developing countries. Bulletin of the World Health Organization 86(4), 282-291.

Veronicah Kirogo, Kogi-Makau Wambui, and Muroki Nelson M. (2007). "The role of irrigation on improvement of nutrition status of young children central Kenya". *African Journal of Food Agriculture Nutrition and Development.* Vol. 7, No. 2; P.1-16.

WoldemariamGirma and Timotiows Genebo (2002).Determinants of the Nutritional Status of Mothers and Children in Ethiopia. Calverton, Maryland, USA: ORC Macro.

World Bank and MoH, (2005). National Strategy for Infant and Young Children, Addis Ababa.

World Health Organization, (WHO, 1995 and 1996). Prevalence of Stunting in child aged 0-4 years related to environmental and socio-economic circumstances. http://www.who.int/ceh/indicators/0_4stunting.pdf

WHO Working Group (1986). Use and interpretation of anthropometric indicators of nutritional status. Bull World Health Organ; **64**:924–41.

WMS, (2004) "Welfare Monitoring Survey: Analytical Report." Statistical Bulletin 339-A, The Federal Democratic Republic of Ethiopia, Central Statistical Authority, Addis Ababa.

Zewditu, G., Kelbessa U., Timotewos, G. and Ayele, N. (2001). Review the Status of Malnutrition and Trend in Ethiopia. Ethiopian Journal of Health and Development 15(2):55-74.

APPENDICES

Appendix A: results of cross-tabulation analysis

Region * Height for age status of the child

Region		Height for age status of the child		Total
		Not stunted	Stunted	
	Tigray	472	429	901
	Affar	405	378	783
	Amhara	511	417	928
	Oromiya	836	501	1337
	Somali	398	151	549
	Benishangul-Gumuz	406	327	733
	SNNP	761	511	1272
	Gambela	438	142	580
	Harari	243	96	339
	Dire Dawa	207	155	362
Total		4677	3107	7784

Pearson chi2(9) = 178.077 Pr = 0.000

Mother's education level * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Mother's education level	No education	2860	1524	4384
	Primary education	1391	1351	2742
	Secondary and above	426	232	658
Total		4677	3107	7784

Pearson chi2(2) = 154.544 Pr = 0.000

Source of drinking water * Height for age status of the child

		Hieght for age status of the child		Total
		Not stunted	Stunted	
Source of drinking water	Non improved source	2751	1681	4432
	Improved source	1926	1426	3352
Total		4677	3107	7784

Pearson chi2(1) = 16.935   Pr = 0.000

Type of toilet facility * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Type of toilet facility	No facilities	3239	1665	4904
	Have facilities	1438	1442	2880
Total		4677	3107	7784

Pearson chi2(1) = 196.533   Pr = 0.000

Number of household members * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Number of household members	1-4	1706	1048	2754
	5-9	2294	1652	3946
	10 and above	677	407	1084
Total		4677	3107	7784

Pearson chi2(2) = 12.772   Pr = 0.000

Wealth index * Height for age status of the child

	Height for age status of the child		Total
	Not stunted	Stunted	

Wealth index	Poor	2594	1898	4492
	Middium	916	587	1503
	Rich	1167	622	1789
Total		4677	3107	7784

Pearson chi2(2) = 30.462   Pr = 0.000

Educational level of Husband/partner's * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Educational level of Husband/partner's	No education	2607	1895	4502
	Primary education	1745	1106	2851
	Secondary and above	325	106	431
Total		4677	3107	7784

Pearson chi2(2) = 52.579   Pr = 0.000

Employment status of Mother * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Employment status of Mother	No	2682	1523	4205
	Yes	1995	1584	3579
Total		4677	3107	7784

Pearson chi2(1) = 52.104   Pr = 0.000

Birth order of the child * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Birth order of the child	1	1480	762	2242
	2-3	1150	938	2088
	4-5	1028	587	1615
	6+	1019	820	1839

Total	4677	3107	7784

Pearson chi2(3) = 80.013 Pr = 0.000

Sex of child * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Sex of child	Male	2192	1577	3769
	Female	2485	1530	4015
Total		4677	3107	7784

Pearson chi2(1) = 11.303 Pr = 0.001

Had diarrhea in last two weeks * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Had diarrhea in last two weeks	No	3972	2577	6549
	Yes	705	530	1235
Total		4677	3107	7784

Pearson chi2(1) = 5.508 Pr = 0.019

Had fever in last two weeks * Height for age status of the child

		Height for age status of the child		Total
		Not stunted	Stunted	
Had fever in last two weeks	No	3774	1213	4987
	Yes	903	1894	2797
Total		4677	3107	7784

Pearson chi2(1) = 1406.853 Pr = 0.000

Age of child * Hieght for age status of the child

		Hieght for age status of the child		Total
		Not stunted	Stunted	
Age of child	0-5 months	1051	172	1223
	6-11 months	782	480	1262
	12-23 months	705	526	1231
	24-35 months	702	583	1285
	36-47 months	745	799	1544
	48-59 months	692	547	1239
Total		4677	3107	7784

Pearson chi2(5) = 462.070 Pr = 0.000

Appendix B: Summary of descriptive statistics for Outliers and Influential observations diagnostics

Descriptive Statistics			
	N	Minimum	Maximum
DFBETA for constant	7784	-.00556	.00869
DFBETA for region	7784	-.00080	.00058
DFBETA for Mother's education	7784	-.00483	.00383
DFBETA for Source of education	7784	-.00490	.00590
DFBETA for types of toilet facilities	7784	-.00601	.00401
DFBETA for Number of household members	7784	-.00271	.00232
DFBETA for Wealth index	7784	-.00159	.00178
DFBETA for Education of husband	7784	-.00303	.00345
DFBETA for Employment status of mother	7784	-.00350	.00302
DFBETA for Birth order of child	7784	-.00143	.00153
DFBETA for Sex of child	7784	-.00140	.00169
DFBETA for had diarrhea in last two weeks	7784	-.00372	.00508
DFBETA for had fever in last two weeks	7784	-.00444	.00193
DFBETA for Age of child	7784	-.00088	.00055
Valid N (listwise)	7784		

Descriptive Statistics			
	N	Minimum	Maximum
Analog of Cook's influence statistics	7784	.00002	.01977
Valid N (listwise)	7784		

Descriptive Statistics			
	N	Minimum	Maximum
Leverage value	7784	.00047	.00603
Valid N (listwise)	7784		