ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCE

DEPARTMENT OF STATISTICS

DETERMINANTS AND TRENDS OF UNDER-FIVE CHILD MORTALITY IN ETHIOPIA

by

Mulugeta Tadesse

A THESIS SUBMITTED TO THE DEPARTMENT OF STATISTICS, ADDIS ABABA UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS (BIOSTATISTICS)

Advisor: Mekonnen Tadesse (Associate Prof.)

JUNE, 2018

ADDIS ABABA, ETHIOPIA

Addis Ababa University

Department of Statistics

This is to certify that the thesis prepared by **Mulugeta Tadesse**, entitled: **Determinants and Trends of Under-Five Child Mortality in Ethiopia: A Multilevel Approach** and submitted in partial fulfillment of the requirements for the Degree of Master of Science in Statistics complies with the regulations of the University and meets the accepted standards with respect to originally 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 declare that the thesis is my original work and all sources and material used for this thesis have been duly acknowledged. This work has not been submitted to any other university for achieving degree or diploma awards.

Name: Mulugeta Tadesse

Signature: ___________________ Date __________________

Place of submission: Department of Statistics, College of Natural and Computational Sciences, Addis Ababa University

Date of Submission: June, 2018

This thesis has been submitted for examination with my approval as University Advisor.

Mekonnen Tadesse (Associate Prof.)	______________________	________________
Advisor's Name	**Signature**	**Date**

ACKNOWLEDGEMENT

First and foremost, I would like to thank God Almighty for giving me the strength, knowledge, ability and opportunity to undertake this research study and to persevere and complete it satisfactorily. Without his blessings, this achievement would not have been possible.

Next I would like to thank Mekonnen Tadesse (Associate Prof.), my thesis advisor, for his tireless support and constructive comments. Without his valuable and encouraging advice, this study would not have been realized.

I would like to thank the Central Statistical Agency for tolerating me until I complete this graduate study successfully.

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

Determinants and Trends of Under-Five Child Mortality in Ethiopia

Mulugeta Tadesse

Addis Ababa University, 2018

ABSTRACT

The burden of under-five mortality remains unevenly distributed. About 80 percent of under-five deaths occur in two regions, sub-Saharan Africa and South Asia. Ethiopia is among the six countries that account for half of the global under-five deaths.The aim of this study was to identify the significant socio-economic and demographic factors influencing under-five child mortality and evaluate the variation among the regional states of Ethiopia.In this study, the 2000, 2005, 2011 and 2016 EDHS data were used to describe the trend of under-five mortality in Ethiopia. The 2016 EDHS data have beenused to analyze determinants and variation of under-five mortality by background characteristics. Single-level logistic regression and multilevel logistic regression models were used to identify the major risk factors of under-five mortality and regional variations in under-five child mortality in Ethiopia using the 2016 EDHS data.The results from single-level and multilevel logistic regression analyses showed that Sex of a child, Age of a child in month, Birth type, Birth order number, Number of Household size, Breastfeeding status, Educational level of mother's, Place of residence and type of toilet facility had significant effects on under-five child mortality and there is variation of under-five child mortality from region to region. Conversely, preceding birth interval, wealth index Household, Source of drinking water and place of delivery were found insignificant.The results revealed variation of under-five child mortality from region to region. The multilevel logistic regression analysis result showed that the effects of breastfeeding varied across regions whereas the effects of other covariates on under-five child mortality remained fixed across regions.

LIST OF ABBREVIATIONS

AIC/ BIC	Akaike/ Bayesian Information Criteria
CSA	Central Statistics Agency
DHS	Demographic and Health Survey
EAs	Enumeration Areas
EDHS	Ethiopia Demographic and Health Survey
EPHI	Ethiopian Public Health Institute
FMoH	Federal Ministry of Health
GTP II	Growth and Transformation Plan II
ICC	Intra-Class Correlation Coefficient
IMR	Infant Mortality Rate
LMIC	Low and Middle Income Countries
MDGs	Millennium Development Goals
MlE	Maximum Likelihood (ML) Estimation
NGO’s	Non-Governmental Organizations
NMR	Neonatal Mortality Rate
EPHC	Ethiopia Population and Housing Census
SDGs	Sustainable Development Goals
SNNP	Southern Nations, Nationalities and Peoples
U5M	Under-Five Mortality
UN-IGME	UN Inter-Agency Group for Child Mortality Estimation
UNICEF	United Nation International Children Emergence Fund

WHO World Health Organization

TABLE OF CONTENT

Title page

ACKNOWLEDGEMENT........III

ABSTRACT........IV

LIST OF ABBREVIATIONS........V

TABLE OF CONTENT........VI

LIST OF TABLES........VIII

CHAPTER ONE........1

INTRODUCTION........1

- **1.1 Backgroundof the Study**........1
- **1.2 Statements of the Problems**........3
- **1.3 Objectives of the Study**........4
- **1.4Significance of the Study**........4

CHAPTER TWO........5

LITERATURE REVIEW........5

CHAPTER THREE........9

DATA AND METHODOLOGY........9

- **3.1 Source of Data**........9
- **3.2Description of the Variables in the Study**........10
  - **3.2.1 The Response Variable**........10
  - **3.2.2 Explanatory Variables**........10
- **3.3 Methods of Data Analysis**........11
  - **3.3.1 Logistic Regression**........12
- **3.4Goodness-of-fit of the model**........13
- **3.5Multilevel Logistic Regression Model**........13

CHAPTER FOUR........15

STATISTICAL DATA ANALYSIS AND RESULTS........15

4.1 Results of Descriptive Analysis 15
4.2 Trends of Under-Five Child Mortality 18
4.3Results of Binary Logistic Regression Analysis 20
4.3.1 Goodness of Fit of the estimated Logistic Regression Model 21
4.3.1.1 The Likelihood Ratio Test 21
4.3.1.2 The Hosmer-Lemeshow Goodness of Fit Test 22
4.3.2 Model diagnostics: influential observations and outliers 23
4.3.3 Interpretation of Binary Logistic Regression model fit results 23
4.4 Results of Multilevel Logistic Regression Analysis 26
4.4.1Results of the Empty Model with Random Intercept 27
4.4.2Results of Random Intercept with Fixed Coefficient Model 30
4.4.3Results of Random Intercept with Random Coefficient Model 32
4.4.4 Multilevel Logistic Regression Model Comparison 35
CHAPTER FIVE **38**
DISCUSSION, CONCLUSION AND RECOMMENDATIONS **38**
5.1 Discussion **38**
5.2Conclusion **41**
5.3 Recommendations **42**
REFERENCES **43**

LIST OF TABLES

Table 4.1: The distribution of U5CM in Ethiopia by socio-economic, demographic and environmental factors: 2016 EDHS data. 16

Table 4. 2: Overall Model Evaluation Using Likelihood Ratio Test (EDHS, 2016). 21

Table 4. 3: Hosmer-Lemeshow Goodness of Fit Statistics 22

Table 4. 4: Binary Logistic Regression model fit results 25

Table 4. 5: Results for Multi-level Logistic Regression Model without covariate 28

Table 4. 6: The Estimate of Random Effect for each Region. 29

Table 4. 7: Estimates for the random intercept with fixed coefficient model 30

Table 4. 8: Results of random intercept with random coefficient model 33

Table 4. 9: Results of multilevel logistic regression model selection criteria 36

LIST OF FIGURES

Figure 4. 1: Trends in childhood mortality rate 19

Figure 4.2: Trends of under-five child mortality by gender 20

Figure 4.3: Trends of under-five child mortality by place of residence 20

Figure 4. 4: Caterpillar plot with the region effects 30

CHAPTER ONE

INTRODUCTION

1.1 Backgroundof the Study

One of the demographic variables that affect population trends/growth is mortality. Infant and child mortality rates are used as summary indicators of social development, quality of life, overall health, child health, maternal health and welfare. Under-five child mortality is a leading indicator of child health and overall development of a nation (Bereka and Habtewold 2017). Child mortality rates also reflect a country's level of the social-economic development and quality of life and are used for monitoring and evaluating population, health programs and policies. Child mortality rates are unacceptably high in many developing countries and need to remain the focus of public policy to gain improvement in infant and child survival.

Every year, millions of children under 5 years of age die, largely from preventable causes (UNICEF, 2017). In almost half of the cases, malnutrition plays a role, while unsafe water, inadequate sanitation and hygiene are also significant contributing factors. The world has made substantial progress in reducing child mortality in the past few decades. Estimates Developed by the UN Inter-agency Group for Child Mortality Estimation (UNICEF, 2017) shows that the global under- five mortality rate declined by 56 percent from 93 deaths per 1,000 live births in 1990 to 41 in 2016. Progress in reducing child mortality has been accelerated during the 2000–2016 period compared with the 1990s. That is, the annual rate of reduction in the under- five mortality rate has increased from 1.9 percent during 1990–2000 to 4.0 percent during 2000–2016. In 2016, an estimated 5.6 million children died before their fifth birthday of which 2.6 million (46 percent) died in the first month of life. It is unacceptable that 15,000 children die every day, mostly from preventable causes and treatable diseases, even though the knowledge and technologies for life-saving interventions are available.

At country level, the under-five mortality rate ranged from a high of 133 deaths per 1,000 live births to a low of 2 deaths per 1,000 live births in 2016. Many countries still have very high rates – about 80 percent of under- five deaths occur in two regions, sub-Saharan Africa and Southern Asia, Six countries account for half of the global under- five deaths, namely, India, Nigeria, Pakistan, the Democratic Republic of the Congo, Ethiopia and China.

Sub-Saharan Africa remains the region with the highest under- five mortality rate in the world. In 2016, the region had an average under- five mortality rate of 79 deaths per 1,000 live births. This translates to 1 child in 13 dying before his or her fifth birthday – 15 times higher than the average ratio of 1 in 189 in high-income countries, or 20 times higher than the ratio of 1 in 250 in the region of Australia and New Zealand. The child health status is always the focus of the health studies in developing countries. It reflects the healthy status and living level and the healthy status in the childhood finally reflects the healthy, income and achievement in the whole life. Children are at greater risk of dying before age five. Reducing child mortality rates has been one of the eight 'Millenniums Development Goals' of the United Nations. Worldwide, under-five child mortality is reducing. But in the developing countries, almost 11 million children are dying every year and two thirds of the deaths are preventable. The unsafe water, malnutrition, the lack of education, health care and social services are the major factors which slow down the reduction rate in some regions. Hence, analysis of child mortality and the methods to reduce the rate are imperative. The high rate countries centralize around Africa, South America and the Southeast Asia (Yifang, 2013).

The 2016 EDHS reported that Ethiopia has successfully reduced the under-five child mortality rate by two-third between 2000 and 2015, which was the target for achieving Millennium Development Goal-4 (CSA, 2016). However, the under-five mortality rate in Ethiopia is still higher than the under-five mortality rates of other several low and middle income countries (LMIC). According to the Ethiopia Demographic and Health Survey (EDHS 2016) result, Neonatal mortality,Infant mortality and Under-5 mortality of Ethiopia declined from 49, 97 and 166 deaths per 1,000 live births in 2000 to 29, 48 and 67 deaths per 1,000 live births in 2016, respectively. This shows that 41%, 50% and 60% decrease in neonatal, infant and under-5 mortality respectively of Ethiopia over a period of 16 years. In other words, in Ethiopia 1 in every 35 children dies within the first month, 1 in every 21 children dies before celebrating the first birthday, and 1 of every 15 children dies before reaching the fifth birthday. Moreover, residence and regional variation are even more pronounced in the mortality of under-five children (CSA, 2016).

On the other hand, thedeterminants of child mortality have not been well investigated using recent data in Ethiopia. The objectives of this study were to study the trends of under-five mortality and identify factors associated with under-five mortality in Ethiopia by using data from the 2016 Ethiopia Demographic and Health Survey.

1.2 Statements of the Problems

Child mortality has received extensive attention internationally through Sustainable Development Goals (SDGs). The burden of under-five deaths remains unevenly distributed. About 80 percent of under-five deaths occur in two regions, sub-Saharan Africa and South Asia. Six countries including Ethiopia account for half of the global under-five deaths(UNICEF, 2017).

For the 5-year period preceding the 2016 survey, the under-five mortality rate was 67 deaths per 1,000 live births which means that 1 in 15 children in Ethiopia dies before reaching age 5 (CSA, 2016). This rate is still high and it is one of the challenging problems that the country needs to address. Most of the studies made on under-five mortality in Ethiopia are based on more than a decade old data though there are few based on the 2011 EDHS which is about seven years old and extrapolations based on outdated surveys may not be reliable for monitoring changes in health status or for comparative analytical work.

As Ethiopia embarks on the implementation of GTP II and Agenda 2030, investing in reliable and timely data to ensure strong monitoring of the new Sustainable Development Goals is crucial. Thus, this study attempted toexplore the major risk factors and assess the regional variation of under-five mortality taking into consideration various health, socio-economic and environmental factors.

In view of the objectives of the study as well as observations made while reviewing relevant literature, the following research questions have been formulated;

- ✓ Which are the factors that significantly influence under-five child mortality in Ethiopia?
- ✓ Are there significant variations in under-five child mortality across the regional states of Ethiopia?
- ✓ What factor(s) have made significant contribution to the variation of under-five child mortality among regional states of Ethiopia?

1.3 Objectives of the Study

General Objective

The main objective of this study was to identify the significant socio-economic and demographic factors influencing under-five child mortality and evaluate the variation of these factors among the regional states of Ethiopia using an appropriate multilevel logistic regression model.

Specific Objectives

The specific objectives of the study which should be accomplished to achieve the general objective stated above are:

- To determine the extent of U5CM within and between regions of Ethiopia.
- To identify the factors that may explain the variation in U5CM between regions of Ethiopia.
- To carry out trend analysis of under-five mortality during 2000-2016.

1.4Significance of the Study

Under-five mortality rate is one of the most important indicators of the socio-economic well-being and public health conditions of a country.Identifying determinants of under-five mortality (U5M) is important for formulating appropriate health programmes and policies that will help to meet the United nations sustainable development goal (SDG) of reducing under-five mortality to at least as low as 25 deaths per 1,000 live births by 2030 (UNICEF, 2017). Itcan also be used to take more cost-effective interventions and policies to reducing child mortality and to improving the health and life expectancy of the society. Specifically, the finding of this study will:

- Help stakeholders in the planning, formulation and implementation of policies concerning the reduction of under-five mortality.
- provide information about the determinants and trends of under-five mortality to stakeholders like the department of probation and child protection, Ministry of Labour and Social Development, Ministry of Health, National council for children, UNICEF, WHO and other NGO's whose beneficiaries are children so as to improve the quality of child care and their health.
- Provide base information to policy makers and researchers that can be used for further studies on under-five child mortality.
- Help stakeholders in making informed decision and plan appropriate interventions.

CHAPTER TWO

LITERATURE REVIEW

Understanding the causes of child mortality provides important public health insights (UNICEF, 2017).

In countries with very high mortality (those with under-five mortality rates of at least 100 deaths per 1,000 live births), approximately half of child deaths are due to infectious diseases. These deaths are largely preventable of the cause. On the other hand, in very-low-mortality countries (those with under-five mortality rates of less than 10 per 1,000 live births), there are almost no under-five deaths from infectious diseases (UNICEF, 2012).

Several studies have been conducted both locally and globally on the determinants of infant and under-five child mortality. Most of the studies have shown significant association between socioeconomic, demographic and environmental factors and child mortality through making use of survey (DHS) or censuses data.

For instance,Mugarura and Kaberuka(2015) used data from the 2006 Demographic and Health Survey in Uganda to examine the factors associated with child mortality. A single level logistic regression and multilevel logistic regression models were fitted to establish the significant factors affecting child mortality in Uganda. By using a random effects model, Sex of a child, duration of breastfeeding, birth weight, education levelof mothers', age of mothers, household wealth were found to be important determinants of child mortality. The results of the Standard Logistic regression model were a bit different from the results of the random effects model. The variables in the random effects model were found to be more statistically significant than those in a standard logistic regression model due to lack of independence of variables in the standard logistic regression model or variation in mortality rates due to hierarchical structure in the data

Bedada(2017) used data from the 2011 Ethiopia Demographic and Health Survey (EDHS, 2011) to examine the existence of regional heterogeneity in under- five children mortality by using single and multilevel binary specifications. The findings of his study revealed that Sex of the child, family size, education level of mothers, age at first birth of mothers, breastfeeding , the use of contraceptive method and regions of children had significant influence on under- five child mortality.

Goro(2007) used data from 1993, 1998, and 2003 DHS surveys in Ghana to examine the determinants of infant and child mortality in three northern regions by using multivariate logistic regression model. The study revealed that education of mothers, birth order of child and marital status of mothers are powerful significant determinants of infant mortality, but only mothers' education had a significant impact on child mortality. Similarly, Twum-Baah(1994) indicated that children born to mothers with higher educational level were associated with lower risk of infant and child mortality as compared to children born to mothers with primary education level or to non-educated mothers.

Senayitand Eshetu(2013) used data from the 2011 EDHS to identify the determinants of under-five child mortality by using Cox hazard proportional model. The study found out that place of residence, mother's age at first birth, mother's age, mother's education and marital status to be statistically significantly associated with child mortality. The result also showed that children living in rural areas face higher risk of mortality than children living in urban areas. This study suggested that children born to unmarried mothers are expected to experience a higher risk of dying than children born to married mothers. In addition, mother's education was found to be a significant determinant of child mortality showing that children born to illiterate mothers experience higher risk of mortality than children born to mothers with primary and higher education. In addition, the study by Van Raalte, et al (2012) found that the mothers' level of education has significant influence on the survival of children.

Solomon and Emmanuel (2016) used data from the 2011 EDHS to identify the determinants of under-five child mortality in high mortality regions of Ethiopia by using multivariable Cox proportional regression models. The results of their study revealed that under-five mortality is significantly associated with preceding birth interval, family size, birth type, breastfeeding status, source of drinking water, and income of mothers. Children born after apreceding birth interval of 2-3 years and 3 years and above were significantly less likely to have died before their fifth birthday than those born within two years. Children who were breastfed, for any period, were 25.5% (HR 1.255, 1.005– 1.567, $p = 0.045$) less likely to have died before their fifth birthday than those who were not breastfed. Increased birth interval time corresponds to a low probability of child mortality.

Desta(2011) used data from 2000 and 2005 DHS to identify the role of socioeconomic, demographic and biological factors of infant and child mortality in Ethiopia by using logistic regression analysis. He found out that marital status and mother's educational levels are statistically significant determinants of infant and child mortality. He also found that children born to mothers older than 20 years of age at first birth are about 55 percent less likely to die before five years as compared to children born to mothers under 15 years of age at first birth. This study also suggested that infant and child mortality widely varied between regions in Ethiopia. However, the likelihood of under-five mortality of children born to mothers of age 15-20 at first birth was not significantly different from those in the reference category.

Lemani (2013) used data from 2004 and 2010 Demography and Health Survey in Malawi to examine the determinants of Infant and Child Mortality in Malawi using logistic regression and cox proportional hazard regression models. In this study, the results indicated that age of mothers and mother's educations were statistically significantly associated with child mortality. Using survival analysis on the 2010 DHS, it was found that the infants of poor and middle income households were more likely to die as compared to those from rich households (OR=2.063 and 1.616, respectively).

Seyoum and Sharma(2014) used data from 2011 Ethiopia Demography and Health Survey to estimate the determinants of under–five mortality in rural parts of Ethiopia using survival analysis and found that mother's educational level was associated with under-five mortality in rural part of Ethiopia.

Chowdhury (2013),based on Bangladesh demographic and health survey (BDHS) conducted in 2007, examined the determinants of under-five mortality in Bangladesh by using Multivariate proportional hazard models. The study found that place of residence; region of residence, mother's age and breastfeeding have significant influence on under-five mortality. This finding showed significantly higher under-five mortality for babies born to residents of rural areas compared to the babies born in urban areas. It showed that the hazard of under-five mortality for children of rural areas is almost 66 percent higher than that for children born and brought up in urban areas.

Kumar and File (2010) used data from 2005 EDHS to examine the selected socioeconomic, bio-demographic and maternal health care factors that determine child mortality in Ethiopia by using employs cross tabulation technique. The results of their studyrevealed that birth intervals with preceding birth have significant impact to lowering the risk of child mortalityamong socioeconomic variables. Their results also confirmed that child mortality risk is associated with birth interval.

CHAPTER THREE

DATA AND METHODOLOGY

3.1 Source of Data

The data for this study have been obtained from fourconsecutive Ethiopia Demography and Health Surveys (EDHSs)conducted in 2000,2005, 2011 and 2016.The 2000, 2005, 2011 and 2016 Ethiopia Demographic and Health Survey (EDHS) wereimplemented by the Central Statistical Agency (CSA). By virtue of its mandate, the CSA has conducted the surveys in collaboration with the Federal Ministry of Health (FMoH) and the Ethiopian Public Health Institute (EPHI) with technical assistance from ICF international, and financial as well as technical support from development partners. All actors in this effort have exerted themselves to get reliable, accurate, and up-to-date data to measure the success of the national development agenda-growth and transformation plan II as well as the sustainable development goals. The 2016 survey was conducted from January 18, 2016, to June 27, 2016, based on a nationally representative sample that provides estimates at the national and regional levels and for urban and rural areas.

The 2000, 2005, 2011 and 2016 Ethiopia Demographic and Health Surveys, were designed to provide estimates for the health and demographic variables of interest for the following domains: Ethiopia as a whole; urban and rural areas (each as a separate domain); and 11 geographic administrative regions (nine regions namely: Tigray, Affar, Amhara, Oromiya, Somali, Benishangul-Gumuz, Southern Nations, Nationalities and Peoples (SNNP), Gambela and Harari regional states and two city administrations :Addis Ababa and Dire Dawa).

Sampling Design of the Survey

The sampling frame used for the 2000, 2005, 2011 and 2016 EDHS data is the Ethiopia Population and Housing Census (EPHC) conducted in 1997(for the first two surveys) and in 2007 (for the latter two surveys) by the Central Statistical Agency (CSA). The 2016 EDHS sample was stratified and selected in two stages. Each region was stratified into urban and rural areas, yielding 21 sampling strata. Samples of EAs were selected independently in each stratum in two stages. In the first stage, a total of 645 EAs (202 in urban areas and 443 in rural areas) were selected with probability proportional to EA size (based on the 2007 EPHC) and with independent selection in each sampling stratum. In the second stage of selection, a fixed number

of 28 households per cluster were selected with an equal probability systematic selection from the newly created household listing. A total of 18,008 households were selected for the sample, of which 17,067 were occupied. Of the occupied households, 16,650 were successfully interviewed, yielding a response rate of 98%.

In the interviewed 16583 households, eligible women aged 15-49 were identified for individual interview; complete interviews were 15683, conducted for yielding a response rate of 95 percent. Information for this study was taken from the birth history section of the Women's Questionnaire. The 2016 EDHS data set has hierarchical structure. The hierarchy for this study follows individuals/child as level-1 and regions as level-2. This means that individuals/children are nested in groups/regions.

In this study, the 2000, 2005, 2011 and 2016 EDHS data are used to describe the trend of under-five mortality in Ethiopia. The 2016 EDHS data is used to analyze determinants and variation ofunder-five mortality by background characteristics.

3.2Description of the Variables in the Study

The independent variables that were considered to influence under-five mortality were selected based on findings of available similar studies and the available data on the subject.

3.2.1 The Response Variable

The response variable for this study is under-five child mortality. Under-five mortality is defined as the probability of dying before completing the fifth birthday. Thus, the outcome variable is the child event before reaching five years of age, which is dichotomous and coded as 1 if the child died in the five years before the survey and 0 if alive.

In short: $Y_i = \begin{cases} 1 \; if \; the \; i^{th} child \; died \; \; before \; five \; years \; of \; age \\ \;\; 0 \; if \; the \; i^{th} \; under - five \; child \; is \; still \; alive \end{cases}$

3.2.2 Explanatory Variables

The explanatory variables included in this study are based on the Mosley and Chen (1984) determinants of childhood morbidity and mortality framework for developing countries, experiences from the available similar studies reviewed above and available data on the subject.

The main predictor variables of under-five child mortalityexplored include Demographic, Socioeconomic and Environmental factors.

The**demographic** variables/factors used in this study include**:-**

- ✓ Sex of child (1=male and 2= female)
- ✓ Age of child in month
- ✓ Type of birth (0=Single birth and 1=Multiple birth)
- ✓ Birth order number (1= first birth, 2= between 2-3, 3= between 4-5 and 4=above 5)
- ✓ Preceding birth interval in month(0= First birth, 1=below 24 months, 2=between 24-47 months and 3= above 47 months)
- ✓ Age of mother at first birth (1=under 20 years , 2= between 20-34 years and 3= 35-49 years)
- ✓ Family size (Number of HH members)(1= between 1-3, 2=between 4-6 and 3= above 6)
- ✓ Breastfeeding status (1=ever breastfeed and 2= never breast feed).

The **socioeconomic variables/factors**include:-

- ✓ Mothers' education level (0=no education, 1=primary education, 2=secondary education and 3= higher education)
- ✓ Wealth index(economic status of HH) (1=Poor, 2=Medium and 3=Rich)
- ✓ Region(1=Tigray,2=Afar,3=Amhara,4=Oromia,5=Somali,6=Benshangule-Gumuz,7=SNNP, 8=Gambella,9=Harari,10=Addis Ababa and 11=Diredawa)
- ✓ Place of residence (1=urban and 2=rural).

And the variables that are classified as **environmental variables** include:-

- ✓ Availability of toilet facility (0=improved facility and 1=unimproved facility and 2=no toilet facility).
- ✓ Source of drinking water (0=protected source and 1=unprotected source)
- ✓ Place of delivery (0=at home and 1=at health center)

3.3 Methods of Data Analysis

In order to study the trend and factors associated with under-five mortality, descriptive statistics, trend analysis, ordinary logistic regression analysis and multilevel logistic regression analysis were used. In the multilevel analysis, three multilevel models (an empty model, model controlling for the individual-level variables, and a model controlling for community-level variables) were constructed.

3.3.1 Logistic Regression

Logistic regression is a popular modeling approach when the dependent variable is dichotomous or polytomous. This model allows one to predict outcomes, from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these. Hosmer and Lemeshow (2000) described logistic regression focusing on its theoretical and applied aspects.

Model description

Binary data are the most common form of categorical data and the most popular model for binary data is logistic regression (Agresti, 2002). Binary logistic regression model is a form of regression, which is used when the dependent variables is dichotomous. In this study, it is used to investigate the effect of predictors on the probability of having under-five child mortality. The dependent variable is given as

$$\mathbf{Y}_{ij}=\begin{cases} 1 \; if \;\; child \; i \; from \; region \; j \; died \;\; before \; five \; years \; of \; age \\ 0 \; if \; the \; under-five \; child \; is \; still \; alive \end{cases} \quad (3.1)$$

i =1, 2, 3…M_j and j = 1, 2, 3 ….N

Where: M-is the number of under-five children in each region j and N-is the number of regions. The conditional probability that the i^{th}child has died given the vector of predictor variables $\mathbf{X}_i$is denoted by$\pi_i = P(y_i = 1|\mathbf{X}_i)$**.** The probability$\pi_i$ in logistic regression model can be expressed in the form of:

$$\pi_i = P(y_i = 1|\mathbf{X}_i) = \frac{\exp(\beta_0+\beta_1 x_{i1}+\cdots\beta_k x_{ik})}{1+\exp(\beta_0+\beta_1 x_{i1}+\cdots\beta_k x_{ik})} = \frac{e^{\mathbf{X}_i{}'\beta}}{1+e^{\mathbf{X}_i{}'\beta}}, i = 1,2, \ldots, n \quad (3.2)$$

Where$\pi_i$in our case is the probability that thei^{th}child dies before five years of age given the vector of predictors ($\mathbf{X}_i$).

The Odds of Ratio

The odds ratio is defined as the ratio of the odds that the event occurs (success) to the odds that the event will not occur (failure). In binary logistic regression analysis, odds ratio is the exponent of the estimated coefficient, $\exp(\hat{\beta})$.

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

$$\log it(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \cdots \beta_k x_{ik} = \sum_{j=0}^{k} \beta_j{}' X_{ij} \qquad (3.3)$$

where , $i = 1,2, \ldots n;\ j = 0,1, \ldots, k$

Each coefficient of a continuous covariate is interpreted as the change in the expected log-odds of having child death before five years of age per unit change of the covariate. In case of categorical predictor variable, it is interpreted as the log-odds of having child death before five years of age with a given category compared to the reference category (Dayton, 1992).

3.4Goodness-of-fit of the model

The goodness of fit of a model measures how effective or well the model describes the response variable. Assessing goodness of fit involves investigating how the predicted values are closer to the observed values. Clearly, the fit is good if there is a good agreement between the fitted and the observed data. Some common approaches to test the goodness of fit of the model are Pearson's χ^2statistic (Hosmer and Lemeshow Test Statistic), Wald test and the likelihood-ratio test (Agresti, 1996).

3.5Multilevel Logistic Regression Model

A multilevel logistic regression model also referred to in the literature as a hierarchical model, can account for lack of independence across levels of nested data (e.g., children nested within regions). Standard logistic regression assumes that all experimental units (in this case, under-five children) are independent in the sense that any variables affecting the dependent variable have the same effect in all regions. Multilevel modeling relaxes this assumption and allows these variables' effects to vary across regions. Because of cost, time and efficiency considerations, stratified multistage samples are the norm for sociological and demographic surveys.For such samples the clustering of the data is, in the phase of data analysis and data reporting, a nuisance which should be taken into consideration.

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. Hence, 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 (Khan and Shaw, 2011).

In this study, multilevel binary logistic regression model would be adopted to uncover the under-five child mortality variations among regional states of Ethiopia. Details of multilevel logistic regression models are provided by Gelman and Hill (2007).

CHAPTER FOUR

STATISTICAL DATA ANALYSIS AND RESULTS

In this chapter, results of our analysis are presented. The 2016 EDHS survey data were analyzed with the help of SPSS version 20 and STATA version 13 statistical (software) packages. The results of our analyses (descriptive, single level binary logistic regression and multilevel logistic regression) are presented in three sections of the current chapter.

4.1 Results of Descriptive Analysis

The distribution of under-five child mortality in Ethiopia corresponding to the socio-economic, demographic and environmental characteristics is presented in Table 4.1. The total number of children aged 0 to 59 month covered in the present study is10, 641.

Of the total number of children included in the study, 48% were female. As displayed in table 4.1, the rate or proportion of under-five child mortality varied from one Region to another with the highest proportion (8.5%) of U5CM recorded in Afar followed by Benishangul-Gumuz (7.3%) and Somali & Harari (6.8%). However, the least proportion of under-five child mortality was observed in Addis Ababa (3.0%) followed by Tigray (4.0%) and Amhara. Hence, there appears to be some variation in the proportion of U5CM among the regions of Ethiopia.

The proportion of under-five child mortality occurred more among males than among females. The prevalence of U5CM among females and males were 5.0% and 6.9%, respectively. The chi-square test results showed a significant association between U5CM and gender (p=0.000).

Under-five child mortality in urban areas was also different from the rate in rural areas. Table 4.1 shows that the proportion of under-five child mortality in urban areas was 3.4% while it was 6.6% in rural areas. This shows that the rate of under-five child mortality in rural areas was about 1.94 times higher than the rate in urban areas. Conversely, the proportion ofunder-five mortality by birth type indicated unexpected variation. A higher percentage (20.9%) of death of under-five children was observed in multiple birth categories as compared to the 5.6% death of under-five children observed in single birth categories.

Table 4.1 shows that the birth order of 6.7% of the children who died was five or more. Similarly, 6.3% of the children that died had first birth order, 5.9% had birth order 4 and 5 while 5.3% were of birth order 2 and 3.

The number of death of under- five children also varied according to household size and age of mothers at first birth. Number of household members indicated unexpected variation of under-five child mortality. Contrary to our expectation, a higher percentage (11.0%) of death of under-five children was observed in households of size three or less and the lowest percentage (4.6%) of death of under-five children was observed in households of size seven or more.On the other hand, a death of about 6.2% was observed among children born to mothers whose age at first birth was under 20 while none of the children born from mothers in the age group 35-49 died.
Breast feeding was important for the survival of under-five children. As expected, the highest percentage (33.2%) of never breast fed under-five children died while mothers who ever breastfed lost only 4.4 % of their children.

The prevalence of death of under-five children varied bymothers' level of education. Table 4.1 reveals that the highest percentage (6.6%) of death of under-five children was observed among mothers' with no education as opposed to the lowest percentage (1.8%) of death of under-five children recorded among mothers with higher education.
Children born to poor families had highest proportion of mortality (6.9%)while children born to rich families had the least proportion of under-five child mortality (4.6%).

The experiences of under-five child mortality were different by the source of drinking water and type of toilet facility. A higher proportion (6.9%) of mortality was recorded among children that drunk water from unprotected sources and a relatively less prevalence (5.3%) of mortality were observed among under-five children that drunk water from protected sources.Similarly, a higher proportion (6.9%) of under-five children died in households with no toilet facility while it was (3.4%) of among children from households having improved toilet facility.

Table4.1: The distribution of U5CM in Ethiopia by socio-economic, demographic and environmental factors: 2016 EDHS data.

Variable	**Under-five child's mortality status**				**Chi-square**	**Df.**	**p-value**
	Alive	**Dead**	**% of U5CM**	**Total**			
Sex of child							
Male	5107	376	**6.9%**	5483	15.970	1	.000**
Female	4899	259	5.0%	5158			
Total	10006	635	6.0%	10641			
Birth type							
single birth	9786	577	5.6%	10363	112.876	1	.000**
multiple birth	220	58	**20.9%**	278			

Birth order number							
First	2031	136	6.3%	2167	5.762	3	.124
between 2-3	3162	176	5.3%	3338			
between4-5	2330	145	5.9%	2475			
above 5	2483	178	**6.7%**	2661			
Preceding birth interval							
first birth	2040	141	6.5%	2181	75.291	3	.000**
below 24	1914	204	**9.6%**	2118			
between 24-47	4157	211	4.8%	4368			
above 47	1895	79	4.0%	1974			
Age of mother at first birth							
below 20	6157	405	**6.2%**	6562	2.528	2	.283
between 20-34	3827	230	5.7%	4057			
between 35-49	22	0	0.0%	22			
Household size							
between 1-3	1156	143	**11.0%**	1299	72.099	2	.000**
between 4-6	4982	304	5.8%	5286			
above 6	3868	188	4.6%	4056			
Breastfeeding status							
Ever breastfeeding,	9622	444	4.4%	10066	804.367	1	.000**
never breastfeed	384	191	**33.2%**	575			
Educational level of mother							
no education	6387	451	**6.6%**	6838	20.699	3	.000**
primary education	2538	140	5.2%	2678			
Secondary	697	37	5.0%	734			
Higher	384	7	1.8%	391			
Wealth index of household							
Poor	5376	399	**6.9%**	5775	21.332	2	.000**
Middle	1386	80	5.5%	1466			
Rich	3244	156	4.6%	3400			
Region							
Tigray	992	41	4.0%	1033	34.474	10	.000**
Afar	972	90	**8.5%**	1062			
Amhara	928	49	5.0%	977			
Oromia	1494	87	5.5%	1581			
Somali	1402	103	6.8%	1505			
Benishangul	815	64	7.3%	879			
SNNPR	1206	71	5.6%	1277			
Gambela	670	44	6.2%	714			
Harari	564	41	6.8%	605			
Addis Ababa	447	14	**3.0%**	461			
Dire Dawa	516	31	5.7%	547			
Type of place of residence							
Urban	1907	67	3.4%	1974	28.602	1	.000**
Rural	8099	568	**6.6%**	8667			
Type of toilet facility							
improved facility	1729	61	3.4%	1790	28.116	2	.000**

unimproved facility	3800	243	6.0%	4043			
no toilet facility	4477	331	**6.9%**	4808			
Source of drinking water							
protected source	6011	338	5.3%	6349	11.627	1	.001**
unprotected source	3995	297	**6.9%**	4292			
Place of delivery							
at home	6781	490	**6.7%**	7271	24.360	1	.000**
at health facility	3225	145	4.3%	3370			

*significant at 5% level, **significant at 1% level.

Test of Association between dependent variable and its independent variables

In order to find out whether each of the independent variables is associated with under-five child mortality, cross tabulations were done and chi-square tests of independence/association were performed.

The null hypothesis for testing the independence of two (categorical) variables is given by: $H_{o:}$ the two (categorical) variables are independent (not associated). If $p < 0.05$, we reject the hypothesis of independence (H_o) at 5% of level of significance.

Using SPSS and STATA with tab or tabulate command, cross-tabulation analysis has been conducted. Based on the resultsdisplayed in Table 4.1, under-five child mortality was found to be associated with sex of a child (p=0.000), Age of child in month (p=0.000), Birth type (p=0.000), Birth order number (p=0.124), Preceding birth interval (p=0.000),household size (p=0.000), breastfeeding status (p=0.000), Educational level of mother's (p=0.000),Wealth index of household (p=0.000), Region (p=0.000),Place of residence (p = 0.000),Type of toilet facility (p=0.000), Source of drinking water (p=0.001) and Place of delivery (p=0.000)at 25% level of significance. That is, the individual contribution of each of these predictors to under-five child mortality is significant. While, the covariate age of mother at first birth was not associated with under-five child mortality at 25% level of significance. Hence, all the significantly associated covariates are considered in the logistic regression model.

4.2 Trends of Under-Five Child Mortality

Before identifying the major determinates of under-five child mortality (U5CM) it is worthwhile to see the trends of U5CM. From 2000, 2005, 2011 to 2016 EDHS data found that all the under-five child mortality rate had declined over the last 16 years. The magnitude of decline varies among the components of under-five child mortality.

The 2016 EDHS showed that the neonatal mortality rate, the infant mortality rate and the under-five child mortality rate were 29, 48 and 67 death per 1000 live births respectively. In other word, in Ethiopia, 1 in every 35 children dies within in the first month, 1 in every 21 children dies before celebrating the first birthday and 1 of every 15 children dies before reaching the fifth birthday.

The neonatal mortality rate declined by 41% from 49 per 1000 live births in the year 2000 to 29 per 1000 live births in 2016.Infant mortality also declined by 50% from 97per 1000 live births in the year 2000 to 48 per 1000 live births in 2016.Similarly under-five mortality has shown a continuous reduction from 166, 123, 88 deaths per 1000 live births in the years2000, 2005, 2011to 67 deaths per 1000 live births in2016. This equates to a decline of 60 percent between the 2000 and the 2016 survey periods (Figure 4.1).

Figure 4. 1: Trends in childhood mortality rate

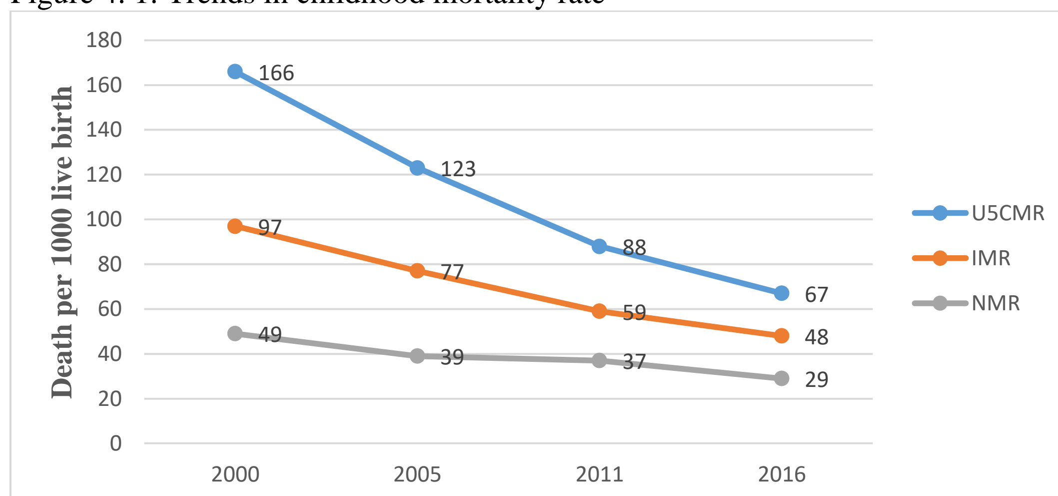


Figure 4.2 shows that U5CM for male children was consistently higher than that for female children between 2000 and 2016. Place of residence and regional variation are even more pronounced in the mortality of under-five children. Over the 16 years period between 2000 & 2016, the under-five child mortality in rural areas was consistently higher than that for urban areas (Figure 4.3).

Generally, the trend analysis U5CM showed that Ethiopia has achieved MDG4.However, U5CM is still higher than the under-five child mortality rate of other several low and middle income countries.

Figure 4.2: Trends of under-five child mortality by gender

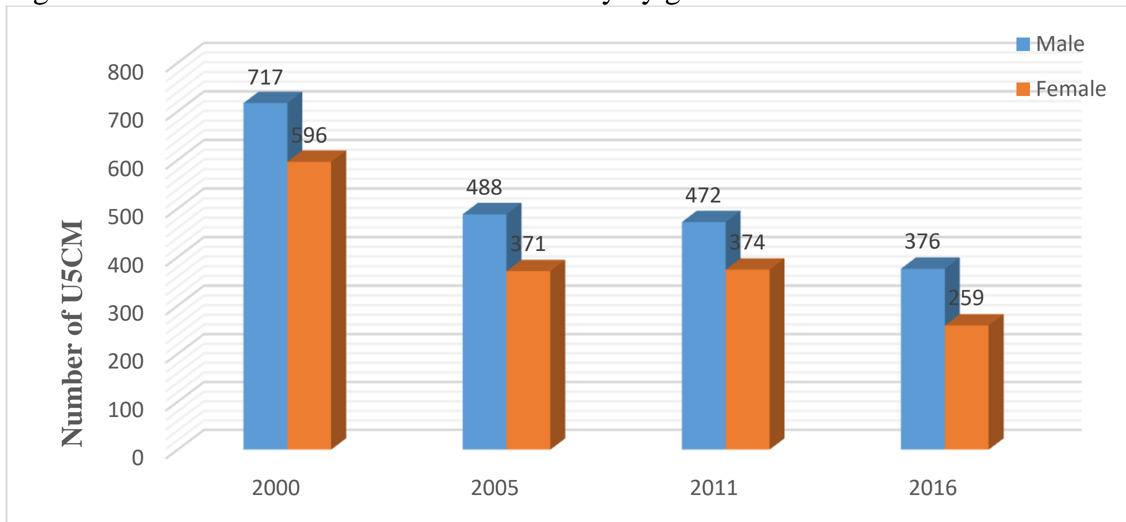
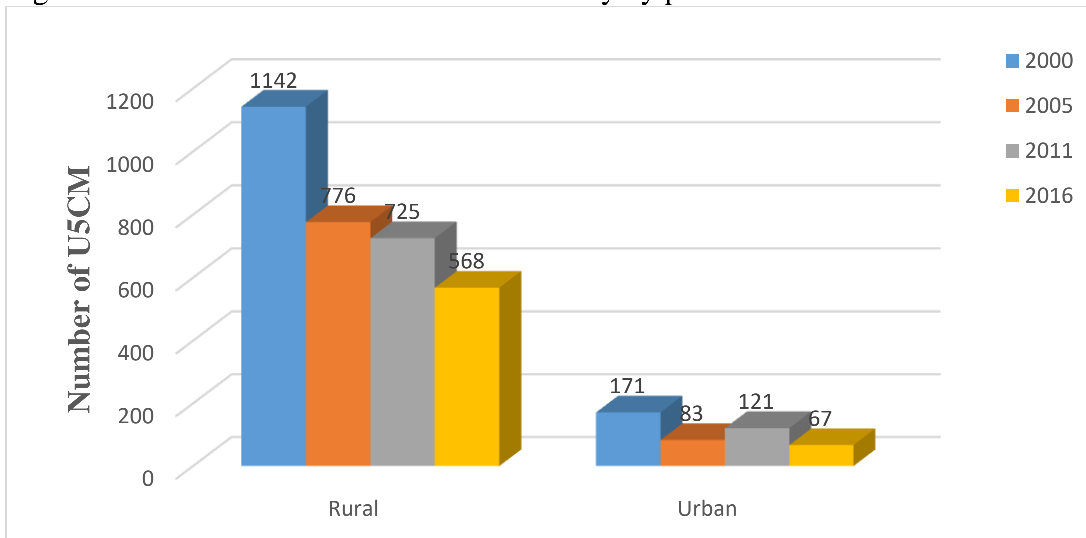


Figure 4.3: Trends of under-five child mortality by place of residence



4.3Results of Binary Logistic Regression Analysis

In this section binary logistic regression was applied to assess the relationship between under-five mortality, which is a dichotomous response variable and the predictor variables.

Multiple binary logistic regression was used to analyze the effects of the explanatory variables on under-five child mortality. In fitting the binary logistic regression model, STATA was used. The statistical significance of the individual regression coefficients have been tested using the Wald chi-square statistic. Accordingly, Sex of the child, Age of the child in month, Birth type, Birth order number, Household Size, breastfeeding status, Region, Place of residence and Type of Toilet Facility were found to be significantly associated with under-five child mortality.

Conversely, Preceding birth interval, Educational level of mothers', wealth index ofHousehold, Source of drinking water and place of delivery were found insignificant at 5% significance level suggesting independence with under-five child mortality (See Table 4.4).

4.3.1 Goodness of Fit of the estimated Logistic Regression Model

Before interpreting the results in Table 4.4, we have to check the goodness of fit of the model to the observed data. In order to check for 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.

4.3.1.1 The Likelihood Ratio Test

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. In our case, we compared two models; one with no variable called the empty (intercept only) model and the other with all variables included called the saturated (full) model. The null and alternative hypotheses to be tested are:

H_0: there is no significant difference between the empty model and the saturated model.
H_A: not H_0

Table 4. 2: Overall Model Evaluation Using Likelihood Ratio Test (EDHS, 2016).

Goodness of Fit measure	Log likelihood(LL)	Deviance =-2LL	Chi-Sq$X^2{}_{0.05}\ (Df)$	Df.	AIC	BIC
Null Model	-2405.631	4811.262	3.84	1	4813.262	4820.535
Full Model	-2009.34	4018.68	43.77	33	4084.68	4324.671

The value for the likelihood ratioG^2statistic is the difference between the -2LL value of the empty model and the -2LL value of the full model. Using the results in table 4.2, we can compute G^2 as the difference in -2LL asG^2= 4811.26-4018.68 =792.58. Since the likelihood ratio test statistics G^2 =792.58 exceeds the tabulated value χ^2 (33) =43.773 and p-value=0.000,we reject the null hypothesis of no significant difference between the two models and conclude that at least one of the predictors'was significantly related with under-five mortality.

To determine the more adequate model, we used AIC and BIC values. The smaller the AIC and BIC values the better the model. Table 4.2 showed that the full model is better than the empty model.

4.3.1.2 The Hosmer-Lemeshow Goodness of Fit Test

The Hosmer-Lemeshow goodness of fit test divides subjects into g classes(often deciles) based on predicted probabilities and then computes a chi-square from observed and expected frequencies (usually in a 10×2 contingency table). A non-significant chi-square indicates that there is no difference between the observed and the model predicted values (classification) and hence estimates of the model adequatelyfit the data.The Hosmer-Lemeshow Goodness-of-fit test tests the hypotheses:

Ho: the model adequately fits the data, vs

Ha: the model not adequately fits the data.

Since the p-value =0.986> 0.05, we do not reject the null hypothesis of no difference between observed and model predicted values, implying that the estimated model adequately fitsthe data (see Table 4.3).

Table 4. 3: Hosmer-Lemeshow Goodness of Fit Statistics

Hosmer-Lemeshow Goodness of Fit Statistics		
Chi-square	Df.	Sig.
2.79	8	.9859

Contingency Table for Hosmer and Lemeshow Test

Group	Prob	child is alive = yes		child is alive = no		Total
		Observed	Expected	Observed	Expected	
1	0.0131	1058	1055.5	7	9.5	1065
2	0.0189	1050	1046.7	14	17.3	1064
3	0.0238	1042	1041.3	22	22.7	1064
4	0.0290	1033	1036.0	31	28.0	1064
5	0.0352	1031	1029.9	33	34.1	1064
6	0.0426	1023	1022.8	41	41.2	1064
7	0.0534	1009	1013.2	55	50.8	1064
8	0.0710	996	998.9	68	65.1	1064
9	0.1111	966	971.2	98	92.8	1064
10	0.8687	798	790.6	266	273.4	1064

Number of observations = 10641 Hosmer-Lemeshow chi2(8) = 2.79

Number of groups = 10 Prob > chi2 = 0.9859

4.3.2 Model diagnostics: influential observations and outliers

So far, we have discussed some summary statistics and examined the goodness of fit of our model. Before concluding that the model is adequate, we have considered measures relevant forthe detection of the presence of outliers and influential observations. The results are presented the appendix (Table A1).

Each one of the model diagnostic measures presented in Table A1 of the appendix is less than unity. DFBETAs less than unity imply no specific impact of an observation on the estimates of the coefficient of a particular predictor variable, while Cook's distance less than unity indicate that an observation had no overall impact on the group estimates of regression coefficients β. Similarly, the value of the leverage statistic is less than one implying that no observation is far apart from the others in terms of the level of thecovariate variables. The fact that the maximum value of Cook's influence statistics and DFBETA for each predictor variable are less than 1.0 indicate that there is no potential influential observation. Based on theabove goodness of fit tests and diagnostic checking results, we can say that our model provide an adequate fit to the data.

4.3.3 Interpretation of Binary Logistic Regression model fit results

The statistical significance of individual regression coefficients is tested using the Wald chi-square statistic. The results in Table4.4 revealed that Sex of the child, Age of the child in month, Birth type, Birth order number, Number of Household Size, Duration of breastfeeding status, Region, Place of residence and type of toilet facility had significant effects on under-five child mortality at 5% level of significance. Conversely, preceding birth interval, Educational level of mother's, wealth index Household, Source of drinking water and place of delivery were found insignificant. The test of significance of the intercept also indicated that it is significant implying that it should be included in the model. For all explanatory variables, the first category was taken as the reference category.

A more appealing way to interpret the regression coefficients in logistic model is using odds ratios. The odds ratio indicates the effect of each explanatory variable directly on the odds of dying (exposure) rather than on the odds of survival (control, unexposed or reference group) of under- five children. Estimates of odds ratio greater than 1.0 indicate that the risk of having under-five child mortality is greater than that for the reference category. Estimates less than 1.0 indicate that the risk of death of an under- five child is less than that for the reference category of

each variablewhile an estimated odds ratio of oneindicate no significant difference in the risk of death between any category and the reference category. So, the standard logistic regression model presented in Table 4.4 is interpreted in terms of odds ratios as follows.

Female children had reduced risks of dying before 5 years of age compared to male children (aOR=0.754, p=0.002). Female children were about 25% (aOR = 0.754, p=0.002) less likely to diebefore 5 years of age compared to male children controlling for other variables in the model.

In this study, child's agehad a statistically significant effect on under-five child mortality. For a one month increase in age, the odds of having mortality increased by 1.3% (OR=1.013).The odds of death of under-five children amongmultiple births was 4.512 (aOR=4.512) times higher than the odds of death of under-five among single births.

Breastfeeding status showed a statistically significant effect on under-five child mortality. Children that were never breastfed had increased risk of dying before five years of age compared to children who were ever breastfed (aOR=11.05).

Children in households of size 4-6 were 67.3% (aOR=0.327) less likely to die before the age of 5compared to children in households of size two or less (below three). Similarly, children in households of size6 or more were 82.9% (aOR=0.171) less likely to die before the age of 5 compared to children in households of size two or less. These results are contrary to our expectations.

When we look at regional effects on under-five child mortality, under-five children in Afar were about 72% more likely to die (aOR=1.719), those in Harari region were about 84% more likely to die (aOR=1.839) and those in Dire Dawa region were about 75% more likely to die (aOR=1.747) compared to under-five children in Tigray region.

Likewise, under-five children in Rural areas were about 69% (aOR=1.692) more likely to die compared to those in Urban areas.

Finally, The odds of under-five child mortality among children in households using unimproved toilet facility were 1.74 times the odds of under-five child mortality among children in households having improved toilet facilities andchildren in households with no toilet facilities

were 1.56times more likely to die compared to under-five children in households with improved toilet facilities, controlling for the other variables in the model.

Table 4. 4:Binary Logistic Regression model fit results

Variables with categories	Coef.	Std. Err.	z	P>z	aOR	95% CI for OR	
SEX OF CHILD(Male= ref.cat)							
Female	-0.283	0.089	-3.16	**0.002****	**0.754**	0.633	0.898
Age of child in month	0.013	0.003	4.69	**0.000****	**1.013**	1.007	1.018
BIRTH TYPE (Single birth=ref.cat)							
multiple birth	1.507	0.187	8.08	**0.000****	**4.512**	3.130	6.504
BIRTH ORDER NUMBER (First birth= ref.cat)							
between 2-3	0.709	0.664	1.07	**0.286**	**2.031**	0.552	7.472
between4-5	1.084	0.678	1.6	**0.110**	**2.958**	0.782	11.182
above 5	1.367	0.683	2	**0.045***	**3.924**	1.029	14.962
PRECEDING BIRTH INTERVAL(first birth= ref.cat)							
below 24	-0.129	0.667	-0.19	**0.847**	**0.879**	0.238	3.247
between 24-47	-0.813	0.665	-1.22	**0.221**	**0.443**	0.120	1.632
above 47	-1.017	0.669	-1.52	**0.128**	**0.362**	0.097	1.342
BREASTFEEDING STATUS(ever breastfeed= ref.cat)							
never breastfeed	2.402	0.111	21.68	**0.000****	**11.050**	8.892	13.731
NUMBERS OF HOUSEHOLD MEMBER (below 3= ref.cat)							
between 4-6	-1.118	0.134	-8.38	**0.000****	**0.327**	0.252	0.425
above 6	-1.767	0.166	-10.66	**0.000****	**0.171**	0.123.	0.236
MOTHER EDUCATIONAL LEVEL(no educate= ref.cat)							
Primary	-0.009	0.119	-0.08	**0.938**	**0.991**	0.784	1.251.
Secondary	0.019	0.220	0.09	**0.93**	**1.020**	0.663	1.568
Higher	-0.813	0.421	-1.93	**0.054**	**0.444**	0.194	1.013
WEALTH INDEX OF HOUSEHOLD poor= ref.cat)							
middle	-0.125	0.151	-0.83	**0.409**	**0.883**	0.657	1.187
rich	0.073	0.147	0.5	**0.618**	**1.076**	0.806	1.436
REGION (Tigray= ref.cat)							
Afar	0.542	0.214	2.53	**0.011***	**1.719**	1.130	2.616
Amhara	0.134	0.232	0.58	**0.565**	**1.143**	0.725	1.801
Oromia	0.082	0.214	0.38	**0.703**	**1.085**	0.713	1.651
Somali	0.209	0.212	0.98	**0.325**	**1.232**	0.813	1.868
Benishangul	0.275	0.230	1.2	**0.231**	**1.317**	0.839	2.067

SNNPR	-0.109	0.223	-0.49	**0.625**	**0.897**	0.580	1.387
Gambela	0.239	0.243	0.98	**0.326**	**1.270**	0.789	2.044
Harari	0.609	0.244	2.49	**0.013***	**1.839**	1.139	2.969
Addis Ababa	0.342	0.376	0.91	**0.362**	**1.408**	0.674	2.939
Dire Dawa	0.558	0.265	2.11	**0.035***	**1.747**	1.039	2.937
PLACE OF RESCIDENCE (Urban= ref.cat)							
Rural	0.526	0.196	2.68	**0.007****	**1.692**	1.152	2.484
TYPE OF TOILET FACILITY (improved facility= ref.cat)							
unimproved facility	0.556	0.181	3.08	**0.002****	**1.743**	1.223	2.484
no toilet facility	0.442	0.185	2.39	**0.017***	**1.556**	1.083	2.236
SOURCE OF DRINKING WATER (protected source= ref.cat)							
unprotected source	0.133	0.097	1.37	**0.17**	**1.142**	0.945	1.382
PLACE OF DELIVERY (at home= ref.cat)							
at health facility	-0.121	0.128	-0.95	**0.343**	**0.886**	0.690	1.138
_cons	-3.704	0.341	-10.87	**0.000***	**0.025**	0.013	0.048

*Significant at 5% level. ** Significant at 1% level, Ref.cat= reference category.

4.4 Results of Multilevel Logistic Regression Analysis

In this study, a two-level structure is used with children at level-1 nested within region at level-2 in order to see the existence of variation with regard to child mortality within and between regions of Ethiopia. In the 2016 EDHS, the 11 regions are considered as level-2 with a total of 10641 children considered as level-1. The data used in this study consist of variables describing individuals as well as groups (regions).

We have considered three multilevel logistic regression models: the empty model, random intercept with fixed effects model and the random coefficient with random intercept model. We have also presented results of model comparison, goodness of fit test and provided interpretations of the fixed effects in terms of odds ratios.

Test of Heterogeneity

Before attempting to multilevel analysis, as one of the aims of this study, we have to test for the heterogeneity of under-five child mortality among regional states of Ethiopia. The chi-square test was applied to assess heterogeneity between regions. As shown in Table 4.1, the Pearson chi-square value is, χ^2 (10)= 34.474with P-value = 0.000, implying that there is strong evidence of heterogeneity of under-five child mortality across regional states of Ethiopia.

4.4.1Results of the Empty Model with Random Intercept

The empty two-level model also called the null two-level model for a dichotomous outcome variable refers to a population of groups (level-two units i.e. regions) and specifies the probability distribution for group-dependent probabilities,π_j. It is the model that incorporates only the grand mean and random intercept (regional effect) without any covariate.

It is given by: - logit(π_j) =β_0 + U_{0j}, where U_{0j} ~IID (0, σ_0^2).

The intercept β_0 also known as the grand (population) 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 under-five children mortality in a particular region deviates from the grand mean. σ_0^2 is the between regions variance.

As shownin Tables 4.2 and 4.9, the deviance based chi-square 10.56= (4811.262-4800.7) is greater than $\chi2$ =3.84 at 1 df and p-value =0.0006. This result indicates that the empty model with random intercept is more appropriate than the empty model without random intercept in predicting under-five child mortality in Ethiopia.

As the results in Table 4.5 show, the estimate of the fixed part of the model is -2.788 which is the log odds of under-five child mortality across all regions. The fixed part of the model is interpreted as the grand mean of the log odds of under-five child mortality with odds of exp(-2.788) =0.062 which is the same as the sample ratio of 635 deaths to 10006 alive.The average probability of under-five child mortality is $\frac{\boldsymbol{exp}^{(-2.788)}}{\mathbf{1}+\boldsymbol{exp}^{(-2.788)}} = \mathbf{0.058}$which means that the chance of under-five child mortality is 5.8% on average. The table alsocontains the variance estimate of the random effects at regional level, σ_0^2=0.044 and we can calculate intra-class correlation coefficient which is the measure of the correlation between two individuals who are in the same higher level unit (region). A low ICC indicates relatively small between region variations. In other words, regions tend to perform at comparable levels to reduce under-five child mortality. As ICC increases, then regions perform with ever increasing variations to reduce under-five child mortality. The between regions variance is0.044whilethe level one variance is $\pi^2/3$ =3.29. From equation $\boldsymbol{ICC} = \frac{\sigma_0^2}{\sigma_0^2+\mathbf{3.29}}$the intra-region correlation has been computed to be 0.013which is very small. This result implies that 1.3% of the variation in the under-five child

mortality can be explained by grouping the children in regions .The remaining (100-1.3%=98.7%) of the variation in the under-five child mortality is explained within region-lower level units.

The random effect tests examine whether or not the random intercept (between-region) variance is needed for these data.The likelihood ratio statistics for testing the null hypothesis,

Ho: $\sigma_0^2 = 0$which means there is no cross-regional variation in U5CM in Ethiopia.

H_1: $\sigma_0^2 > 0$. For this hypothesis, the variance component test lies on the boundary of the parameter space, the likelihood ratio test can break down asymptotically. At the bottom of the table, the value of the test statistics and the corresponding p-value for testing the hypothesis H_0: σ_0^2=0 that there is no cross-regional variation in under-five mortality are presented. Since the value of the test statistic is 10.56 with p = 0.0006, the null hypothesis is rejected and we conclude that there is strong evidence of heterogeneity or cross-regional variation in under-five mortality incidence in Ethiopia.

Table 4. 5: Results for Multi-level Logistic Regression Model without covariate

Fixed –effect	Coef.	Std. Err.	Z	P>z	[95% Conf. Interval]	
_cons	-2.788	0.077	-35.98	0.000	-2.940	-2.636
Random-effects parameters	Estimate	Std. Err.	[95% Conf. Interval]			
REGION: Identity Var (Uoj) = σ_0^2	0.044	0.029	0.012		0.162	
Rho (ρ)=ICC	0.013					

LR test vs. logistic regression: chibar2(01) = 10.56 Prob>=chibar2 = 0.0006, log likelihood = -2400.35, deviance= 4800.7, AIC= 4804.7 and BIC= 4819.245 with 2 df.

We can now write the model for the jth region as logit(π_j) =-2.788 +U_{0j}. Based on the estimated model, we can say that the average probability of under-five child mortality 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.

Table 4.6 contains the estimated values of U_{0j}. The results indicate that the probability of under-five child mortality is less than the average in Addis Ababa, Tigray, Amhara, Oromia, SNNPR and Dire Dawa while it is higher than the average for the remaining regions. The worst situation has been observed in Affar and Benishangul-Gumuz.

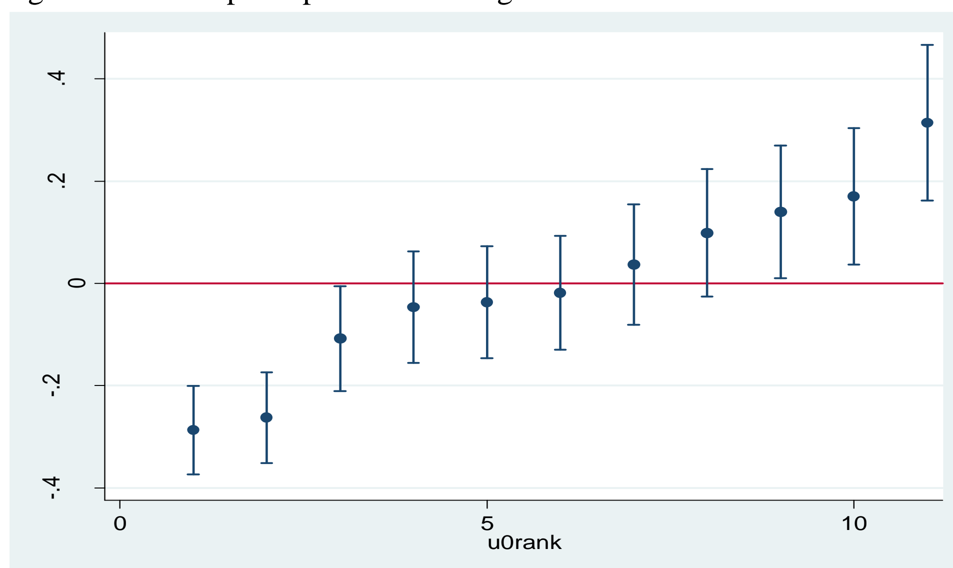
Table 4. 6: The Estimate of Random Effect for each Region.

Rank	Region	Uoj	Standard error Uoj
1	Addis Ababa	-0.2870136	0.0441555
2	Tigray	-0.2623423	0.0452085
3	Amhara	-0.1078248	0.052367
4	Oromia	-0.0467102	0.0554841
5	SNNPR	-0.0368982	0.0560006
6	Dire Dawa	-0.0183893	0.0569871
7	Gamebela	0.03371038	0.0600437
8	Harari	0.0985048	0.0636041
9	Somali	0.139467	0.0660878
10	Benishangul	0.1704585	0.0680265
11	Afar	0.3136441	0.0776853

The Caterpillar Plot with the Region Effects with 95% Confidence Intervals

The caterpillar plot (Fig 4.4), with the region effects shown in rank order (together with 95% confidence interval)shows the estimated region effects or residuals for all 11 regions in the sample obtained from the null model. For the majority of the regions, the 95% confidence interval does not overlap the horizontal line at zero, indicating that uptake of under-five child mortality in these regions was significantly above average (above the zero line) or below average (below the zero line). This indicates that the intercept of under-five child mortality varied across the regions.

Figure 4. 4: Caterpillar plot with the region effects



4.4.2Results of Random Intercept with Fixed Coefficient Model

In a random intercept with fixed coefficient model, we allow the probability of under-five child mortality to vary across regions assuming that the effects of the explanatory variables are the same for each region. That is, the random intercept varies across regions, but children level covariates variables are considered fixed across regions. The results of the two-level random intercept with fixed coefficient model obtained using STATA are presented in the following tables.

Table 4. 7: Estimates for the random intercept with fixed coefficient model

COVARIATES	Coef.	Std.Err.	Z-value	**p>z**	**aOR**	95% Conf. Interval for OR	
SEX OF CHILD(Male= ref.cat)							
Female	-0.282	0.089	-3.17	**0.002****	**0.754**	0.633	0.898
AGEOF CHILDS (M)	0.012	0.003	4.61	**0.000****	**1.013**	1.007	1.018
BIRTH TYPE (Single birth=ref.cat)							
Multiple birth	1.508	0.186	8.1	**0.000****	**4.519**	3.138	6.508
BIRTH ORDER NUMBER (First birth= ref.cat)							
between 2-3	0.738	0.670	1.1	**0.27**	**2.091**	0.563	7.763
between4-5	1.104	0.683	1.62	**0.106**	**3.0175**	0.791	11.511
above 5	1.379	0.687	2.01	**0.045***	**3.971**	1.032	15.276
PRECEDING BIRTH INTERVAL(first birth= ref.cat)							
below 24	-0.149	0.671	-0.22	**0.824**	**0.862**	0.231	3.211
between 24-47	-0.852	0.669	-1.27	**0.203**	**0.427**	0.115	1.585

above 47	-1.071	0.673	-1.59	**0.112**	**0.343**	0.0923	1.282
BREASTFEEDING STATUS(ever breastfeed= ref.cat)							
never breastfeed	2.382	0.110	21.61	**0.000****	**10.828**	8.724	13.440
HOUSEHOLD SIZE (below 3= ref.cat)							
between 4-6	-1.118	0.133	-8.41	**0.000****	**0.327**	0.252	0.424
above 6	-1.762	0.165	-10.65	**0.000****	**0.172**	0.124	0.238
MOTHER EDUCATIONAL LEVEL(no educate= ref.cat)							
Primary	-0.033	0.118	-0.28	**0.78**	**0.968**	0.768	1.219
Secondary	-0.007	0.216	-0.03	**0.976**	**0.993**	0.651	1.516
Higher	-0.845	0.420	-2.01	**0.044***	**0.429**	0.189	0.978
WEALTH INDEX OF HOUSEHOLD poor= ref.cat)							
Middle	-0.159	0.149	-1.07	**0.287**	**0.853**	0.636	1.143
Rich	0.052	0.144	0.36	**0.716**	**1.054**	0.794	1.398
PLACE OF RESCIDENCE (Urban= ref.cat)							
Rural	0.453	0.180	2.39	**0.017***	**1.573**	1.085	2.281
TYPE OF TOILET FACILITY (improved facility= ref.cat)							
unimproved facility	0.502	0.176	2.85	**0.004****	**1.652**	1.170	2.334
no toilet facility	0.425	0.181	2.35	**0.019***	**1.530**	1.073	2.182
SOURCE OF DRINKING WATER (protected source= ref.cat)							
unprotected source	0.127135	0.095778	1.33	**0.184**	**1.13557**	0.9412132	1.370
PLACE OF DELIVERY (at home= ref.cat)							
at health facility	-0.122	0.125	-0.97	**0.331**	**0.885**	0.693	1.132
_cons	-3.334	0.281	-11.85	**0.000****	**0.036**	0.021	0.062
Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]				
REGION: Identity Var(_cons)	0.023	0.020	0.063	0.361			

*significant at 5% level,

LR test vs. logistic model: chibar2(01) = 3.21 Prob >= chibar2 = 0.0367 log likelihood=-2018.79, deviance= 4037.58, AIC= 4085.58 and BIC= 4260.12 with 24 df.

The results for the fixed part of the random intercept with fixed coefficient model show that the sex of child, age of child, birth type, birth order number (above five), breast feeding, household size, mothers' educational level (higher), place of residence and type of toilet facility are significant determinants of variation in under-five child mortality of all regions with respect to the corresponding reference categories (see Table 4.7). The estimated coefficients and odds ratios have similar interpretation as in binary logistic regression discussed above. However, the result for the random part has additional information which is discussed below.

The results in Table 4.5 and Table 4.7show that, the variance component representing variation between regions has decreased from 0.044 in the empty model with random intercept to 0.023 in the random intercept with fixed coefficients multilevel logistic regression model. The reduction of the random effect of the intercept variation is due to the inclusion of fixed explanatory variables. That is, taking into account the fixed independent variables can provide extra prediction values on under-five child mortality in each region.

The deviance-based chi-square,G^2 =18.9 is the difference in deviance between the full model for the single level logistic regression model (deviance=4018.68) and the random intercept with fixed coefficient model (deviance=4037.58). The likelihood ratio test of the null hypothesis that there is no significant difference between the two models gives G^2= 18.9. This value is compared to a chi-square distribution with 9 degrees of freedom. The tabulated value was $\chi2$ (9) =16.91 and,G^2= 18.9> $\chi2$ (9) =16.91, with p = 0.000<0.05. This implies that the null hypothesis should be rejected which indicates that the random intercept with fixed coefficients model is a better fit as compared to the empty model with random intercept. In addition, the AIC and BIC values for the random intercept with fixed effect model (AIC=4085.58, BIC=4260.12) are less than that of the intercept-only model with random effect (AIC=4804.7, BIC = 4819.245). These results indicate that the random intercept with fixed coefficient model gives a better fit to the data than the empty model with random intercept for predicting under-five child mortality among regions in Ethiopia (see Table 4.9). In addition to this, the test of the null hypothesis Ho: $\sigma_0^2 = 0$ that there is no cross-regional variation in under-five children mortality in Ethiopia is rejected because the likelihood ratio test versus logistic regression resulted in chibar2 (01) = 18.9 Prob≥ chibar2 = 0.000 implying evidence of heterogeneity or cross-regional variation in mortality of under- five children for the random intercept with fixed effects model. We can, therefore, conclude that the random effect at regional level is significantly different from zero.

4.4.3Results of Random Intercept with Random Coefficient Model

Multilevel logistic regression analysis allows the coefficient of level-one independent variables to vary across regions instead of keeping them fixed across the regions. This allows region to have different coefficient, implying that the coefficient of independent variables are random at level two (region-level). The effect of breastfeeding has been examined by allowing it to vary

randomly across regions. We investigated whether level-one independent variables have random effects across regions or the same effects across regions.

Estimates of this model show that the variance of the random coefficient of all included variables is zero except that for breastfeeding. This indicate that only the effect of breastfeeding on under-five child mortality varied across regions whereas the effect of the other independent variables on under-five child mortality remain fixed across regions. Results of the random coefficient estimates are presented in Table 4.8 below.

Table 4. 8: Results of random intercept with random coefficient model

COVARIATES	Coeff.	Std.Err.	Z	p>z	aOR	95% Conf. interval for OR	
SEX OF CHILD(Male= ref.cat)							
Female	-0.268	0.090	-2.98	0.003**	0.765	0.641	0.912
AGEOF CHILDS (M)	0.013	0.003	4.77	0.000**	1.013	1.008	1.019
BIRTH TYPE (Single birth=ref.cat)							
Multiple birth	1.520	0.189	8.06	0.000**	4.574	3.161	6.620
BIRTH ORDER NUMBER (First birth= ref.cat)							
between 2-3	0.817	0.675	1.21	0.226	2.265	0.603	8.499
between4-5	1.193	0.689	1.73	0.084	3.296	0.854	12.721
above 5	1.458	0.693	2.1	0.035*	4.299	1.105	16.726
PRECEDING BIRTH INTERVAL(first birth= ref.cat)							
below 24	-0.254	0.677	-0.37	0.708	0.776	0.206	2.925
between 24-47	-0.920	0.675	-1.36	0.173	0.399	0.106	1.496
above 47	-1.193	0.679	-1.76	0.079	0.303	0.080	1.148
BREASTFEEDING (ever breastfeed= ref.cat)							
never breastfeed	2.584	0.324	7.98	0.000**	13.253	7.025	25.002
HOUSEHOLD SIZE (below 3= ref.cat)							
between 4-6	-1.140	0.134	-8.49	0.000**	0.320	0.246	0.416
above 6	-1.795	0.167	-10.76	0.000**	0.166	0.120	0.230
MOTHER EDUCATIONAL LEVEL(no educate= ref.cat)							
Primary	-0.039	0.118	-0.33	0.744	0.962	0.763	1.213
Secondary	-0.020	0.217	-0.09	0.926	0.980	0.640	1.501
Higher	-0.875	0.432	-2.02	0.043*	0.417	0.176	0.973
WEALTH INDEX OF HOUSEHOLD poor= ref.cat)							
Middle	-0.127	0.151	-0.84	0.402	0.881	0.655	1.185

Rich	0.078	0.145	0.54	0.588	1.081	0.815	1.436
PLACE OF RESCIDENCE (Urban= ref.cat)							
Rural	0.498	0.195	2.56	0.011*	1.645	1.123	2.408
TYPE OF TOILET FACILITY (improved facility= ref.cat)							
unimproved facility	0.477	0.179	2.67	0.008**	1.611	1.135	2.286
no toilet facility	0.444	0.184	2.41	0.016*	1.559	1.087	2.237
SOURCE OF DRINKING WATER (protected source= ref.cat)							
unprotected source	0.117	0.096	1.22	0.221	1.125	0.932	1.357
PLACE OF DELIVERY (at home= ref.cat)							
at health facility	-0.151	0.128	-1.18	0.237	0.860	0.669	1.105
_cons	-3.428	0.289	-11.86	0.000**	0.032	0.018	0.057
REGION							
var(BFEED)=Var(U_{6j})	0.967	0.495			0.967	0.354	2.639
var(cons) =Var(U_{0j})	1.451	0.743			1.451	0.532	3.960
Cov (U_{0j},U_{6j})) = σ_{06}	-1.182	0.604	-1.96	0.05	-1.182	-2.366	0.002

LR test vs. logistic regression: chi2(3) = 46.02 Prob > chi2 = 0.0000
Log likelihood=-1997.39, deviance=3994.78, AIC=4046.77 and BIC=4235.86 with 26 df.

In Table 4.8, the values of var(U_{0j}) and var(U_{6j}) are the estimated variance of the intercept and the coefficient of breastfeeding respectively. These estimated variances are significant which suggest that the intercept and coefficient of breastfeeding status vary significantly. So, there is a significant variation in the effect of breastfeeding across regionsof Ethiopia.

The effect of the intercept on region j is estimated to be -3.428+U_{0j}. The intercept variance of 1.45 with standard error 0.7is interpreted as the between-region variance when all other variables are held constant (i.e. equal to zero). Their mean is -3.428 with standard error 0.289. The between –region variance for the coefficient of breastfeeding is estimated to be 0.967 with standard error is 0.495. The negative covariance estimate of -1.182 with standard error 0.604 between intercept and coefficient of breastfeeding suggest that regions with a higher intercept than average tend to have a flatter-than average coefficient.

Generally, interpretation of significant covariance terms can be easily made in terms of the correlation coefficient between the random intercept and the random coefficient. Positive covariance/correlation between the random intercept and the random coefficient implies that regions with higher intercepts tend to have on average higher coefficient on the corresponding predictors. The negative sign for the correlation between the random intercepts and coefficients

imply that regions with higher intercepts tend to have on average lower coefficients on the corresponding predictors. The intercept-coefficient correlation between intercept and coefficient of breastfeeding is estimated as follows.

$\rho_{02}=\frac{-1.182}{\sqrt{0.967*1.451}}$= -0.998 (correlation between intercept and coefficient of breastfeeding).

The random coefficient logistic regression model involves two extra parameters: the variance of the coefficient residuals (i.e. breastfeeding status), U_{6j} and their covariance with the intercept residuals σ_{06} . The change (which is also the change in deviance) can be regarded as a χ^2 value with 1 degree of freedom. Under the null hypothesis that the extra parameters have population value of zero, the value of the deviance based chi-square is given by (4037.58-3994.78=42.8, p-value = 0.000) which shows that the addition of this fixed effect and one random coefficient has significantly improved the fit of the more elaborate model to the data.

The parameter estimates of the observed variables can be interpreted much the same way as those from the standard logistic regression model. For instance, everything else being equal except slight differences on the random effect in the model, under-five children born to mothers' with higher education were 58.3% lesslikely to die (OR=0.417) compared to under-five children born to mothers' with no education controlling for other variables in the model and random effect at level two.

4.4.4 Multilevel Logistic Regression Model Comparison

The deviance, likelihood ratio test, AIC and BIC values were used for selecting the best fitting model among the three fitted two level logistic regression models considered. Table 4.9 shows that the deviance of the empty model with random intercept (4800.7) is greater than the deviance of the random intercept with fixed coefficient (4037.58) and also the deviance of the random intercept with fixed coefficient (4037.58) is greater than the deviance of the random coefficient model (3994.78). These indicate that the random intercept with random coefficient model is better than the multilevel empty model and also the random coefficient model is better than the random intercept with fixed coefficient model.

Similarly, the values of the Akaike Information Criterion (AIC) were used to make an overall comparison of the three models. Table 4.9 also shows that the AIC value for the random

coefficient model is less than that of the random intercept with fixed coefficients model and the empty model with random intercept. This indicates that the random coefficient model provides a better fit as compared to the empty model with random intercept and the random intercept with fixed effect model.

Table 4. 9: Results of multilevel logistic regression model selection criteria

Model selection criteria	Log likelihood(LL)	Deviance -2LL	Deviance based on chi-square	p-value	Df.	AIC	BIC
Multilevel Empty model	-2400.35	4800.7	10.56	0.0006	2	4804.7	4819.25
Multilevel Random intercept model	-2018.791	4037.58	18.9	0.000	24	4085.58	4260.12
Multilevel Random coefficient model	-1997.39	3994.78	42.8	0.000	26	4046.77	4235.86

Goodness of fit test

An overall evaluation of the multilevel logistic regression model was assessed using deviance. The test is done by comparing the deviance (-2 log likelihood) of two models by subtracting the smaller deviance (model with smaller parameters) from the larger deviance (model with greater parameters). The difference is a chi-square test with degrees of freedom equal to the difference inthe number of parameters of the two models. The chi-square test results indicate that the model is a good fit. Similarly, the overall model evaluation was assessed using AIC and BIC. Based on the results in Table 4.9, the Random coefficient model is a good fit.

The likelihood ratio test of no significant difference between the random intercept with fixed effect and random coefficient model gives LR=42.8 (which is the difference between the deviance of the random intercept with fixed effect (4037.58) and the random coefficient (3994.78) with p= 0.000. This implies that there is a significant difference between the two nested models. Similarly, the values of fit statistics for the random coefficient logistic regression model (AIC= 4046.77 and BIC=4235.86) are less than those for the random intercept with fixed coefficient model (AIC= 4085.583 and BIC= 4260.122). These indicate that the random coefficient logistic regression model provides a better fit to the data than the random intercept

with fixed effect model. Thus, all model comparison criteria revealed that the random coefficient modelis the best fitting model among the three two-level models considered.

CHAPTER FIVE

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

In this chapter, we discuss the findings of this study and also draw conclusions and forward recommendations based on the results obtained.

5.1 Discussion

This study has attempted to identify the significant socio-economic and demographic factors influencing under-five child mortality and evaluate the variation of these factors among the regional states of Ethiopia using the Ethiopian Demographic and Health Survey (EDHS, 2016) data set. Descriptive statistics, Single-level logistic regression and multilevel logistic regression statistical methods of data analyses were employed to identify determinants of under-five mortality in Ethiopia. The obtained results are discussed as follows.

The results of this study showed that children in Afar and Benishangul-Gumuz regionswere at a higher risk of death than children in Tigray. This study showed that region was a significant factor and some variation of U5CM was observed between regions.Under-five children in Afar, Harari and Dire Dawa were more likely to die compared to under-five children in Tigray region. This disparity could be due to differences in the distribution of socio economic, infrastructure and functioning of the healthcare system in each region. This result is consistent with the finding of Bedada (2017) in Ethiopia.

The results of this study showed that Female children had reduced risks of dying before 5 years of age compared to male children. Female children were 25% less likely to experience under-five deaths compared with male children. This finding is complementary to the findings ofBedada(2017) and Mugarura and Kaberuka (2015). They found out that the risk of death of under- five female children was less than the risk of death of male under-five children.

The results of the present study indicate that age of a child is one of the important determinant factors of under-five child mortality in Ethiopia. For a one month increase in age, the odds of having mortality increased by 1.3% (OR=1.013).

The results of this study have revealed that births type has a statistically significant effect on U5CM.Multiple births were more associated with under-five child mortality than single births.

This result is consistent with the results of the study by Bedada (2017). Similarly, breastfeeding status showed a statistically significant effect on U5CM. Never breastfed under-five childrenuntil the survey time were more likely to die than ever breasted under-five children. This result is consistent with the results of the study by Bedada (2017) and the study by Solomon and Emmanuel (2016) in Ethiopia. They showed that there is a higher under-five child death among children who were not breastfed than those ever breastfed fed.

Birth order of a child showed a statistically significant effect on under-five child mortality. Some studies showed that first births were at a higher risk of under-five child mortality while others showed that higher ranked births were at increased risk of under-five child mortality. For instance, a study conducted in Taiwan (Wang and Lin, 1999) showed that children with first and fifth ranked births were at higher risk of early under-five child mortality. Hailemariam and Tesfaye (1997) showed that children with sixth or higher order birth were at increased risk of under-five child mortality in Ethiopia. Our study showed that five or more ranked births were at an increased risk of under-five child mortality compared to first ranked births.

Education plays a significant role in reducing under-five child mortality, particularly maternal education has a greater contribution towards the reduction of under-five child mortality. In line with this, Senayit and Eshetu (2013) found that child mortality in Ethiopia is highly associated with mother's education. The results showed that children born to mothers' with no education experience higher risk of mortality than children born to mothers with primary and higher education. Similarly, Kumar and File (2010) in Ethiopia found similar results that mothers' education and under-five child mortality are significantly associated. In this study, it was found that children born to mothers with higher education were less exposed to mortality than children born to mothers with no education.

Household size plays an important role in under-five child mortality. The results of this study revealed that household size has a significant effect on U5CM:Children in households of size 4 and above were less likely to die before the age of 5 compared to children in households of size two or less (below three). This result is consistent with the results of the study by Desta that was carried out in Ethiopia. His results indicated an inverse relationship between under-five child mortality and numbers of household members and contrarily, the study by Solomon & Emmanuel

carried out in Ethiopia found direct relationship between under-five child mortality and number of household members. These results are contrary to our expectation. This discrepancy could be due to the time gap between the current study and the study by Desta (2011) and Solomon and Emmanuel (2016).

Place of residence is significantlyassociated with under-five child mortality. Children born in rural areas were more likely to die than those born in urban areas. This result is consistent with the findings of Senayit and Eshetu (2013) that children living in rural areas face higher risk of mortality than children living in urban areas. This result also supports the earlier findings of Chowdhury (2013)which showed that the risk of under-five mortality for children born in rural areas is almost 66 percent higher than the risk for children born in urban areas.

Finally, this study found out thattype of toilet facilities and U5CM are associated. The odds of under-five child mortality among children in households using unimproved toilet facility were more than the odds of under-five child mortality among children in households having improved toilet facilities and children in households with no toilet facilities were much more likely to die compared to under-five children in households with improved toilet facilities, controlling for the other variables in the model.In line with this, Chowdhury (2013) found that the risk of under-five mortality is higher for children living in household without any toilet compared to the children living in houses with improved toilet facilities.

5.2Conclusion

This study was intended to identify some determinants of under-five child mortality in Ethiopia based on Ethiopia Demographic and Health Survey (EDHS 2016) data. Accordingly, descriptive analysis, single level logistic regression and multi-level logistic regression analyses techniques were used.

The study included the most important predictor variables that were categorized under socio-economic, demographic and environmental characteristics. The results of this study revealed that Sex of a child, Age of child in months, Birth type, Birth order number, Household size, breastfeeding status, Mothers level of education, Region, Place of residence and type of toilet facility were among the determinants of under-five child mortality in Ethiopia.

The random coefficient model was found to be the best fitting model among the three two-level models considered. Under-five child mortality variation among regional states is accounted by the random intercept of the model. Moreover, the variance of the random component, related to the intercept term, is found to be significant implying the presence of under-five child mortality variation across regional states.

The variation between regions is very small as shown by the findings of this study. This shows that regions tend to perform at comparable levels towards reducing under-five child mortality. However, there is a variation across regional state of Ethiopia up until now.

According to the results of the multilevel logistic regression analysis, the effects of breastfeeding varied across regions whereas the effects of the other covariates on under-five child mortality remained fixed across regions.Thus, we conclude that breastfeeding had significant impacts on under-five child mortality and varies across regional states of Ethiopia. The analysis of the final model indicated significant regional-level variation. This may suggest differences in lifestyles, culture, or environment between different regions. Because of these potential cultural, socio-economic and environmental differences, under-five child mortality exhibits a significant variation among regions of Ethiopia.

5.3 Recommendations

Based on the results of the study, the following recommendations can be made.

- To reducemore under-five deaths and enhance child survival, efforts by the government and other stakeholders in the health sector must be directed at factors identified by the study.
- The results of our study revealed that multiple births are associated with under-five child mortality. This calls for the improvement of maternal and child health services that contribute to improvement of child survival rates.
- The severity of U5CM varies from one region to another. Future studies should focus on identifying the risk factors of U5CM for each region of Ethiopia separatelyin order to inform policy makers to formulate region specific strategy. The government should give more attention to those regions with high under-five mortality rates (like Affar and Benishangul-Gumuz) so that the rate in these regions is substantially reduced.
- Breastfeeding plays an important role in child survival. Government and other concerned bodies should give more attention to creating more awareness about the benefits of breastfeeding.
- Finally, we recommended that further research should be conducted to identify the determinants of under-five child mortality in Ethiopia including other factors that were not included in this study due to absence of recorded data.

REFERENCES

Agresti, A. (1996).*An Introduction to Categorical data Analysis*: John Wiley and Sons, Inc., New York.

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

Bedada, D. (2017). Determinant of Under-Five Child Mortality in Ethiopia. *American Journal of Statistics and Probability*, *2*(2), 12-18.

Bereka, S. G., & Habtewold, F. G. (2017). Under-Five Mortality of Children and its Determinants in Ethiopian Somali Regional State, Eastern Ethiopia. *Health Science Journal*, *11*(3).

Central Statistical Agency (CSA.2016) [Ethiopia] and ICF International. 2016. *Ethiopia Demographic and HealthSurvey 2016*. Addis Ababa, Ethiopia and Rockville, Calverton, Maryland, USA: Central Statistical Agency and ICF International.

Chowdhury, A. H. (2013). Determinants of under-five mortality in Bangladesh. *Open Journal of Statistics*, *3*(03), 213.

Dayton, C. M. (1992). Logistic Regression Analysis; Department of Measurement. *Statistics and Evaluation, University of Maryland*.

Desta, M. (2011). Infant and Child Mortality in Ethiopia. *The role of Socioeconomic, Demographic and Biological factors in the previous five years period of 2000 and 2005*.

Ethiopia Demographic and Health Survey (EDHS). Central Statistics Authority & ORC Marco. Addis Ababa, Ethiopia and Calverton, Maryland, USA, 2016.

Gelman, A, &Hill, J (2007). Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.

Goro, M. (2007). The stalling child mortality: the case of three northern regions. In *The 5th conference of union for Africa population, Tanzania*.

Hailemariam, A., & Tesfaye, M. (1997). Determinants of infant and early childhood mortality in a small urban community of Ethiopia: a hazard model analysis. *Ethiopian Journal of Health Development*, *11*(3), 189-200.

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

Khan, H. R., & Shaw, E. (2011). Multilevel logistic regression analysis applied to binary contraceptive prevalence data.

Kumar, P. P., &File, G. (2010). Infant and child mortality in Ethiopia: a statistical analysis approach. *Ethiopian Journal of Education and Sciences, 5(2).*

Lemani, C. (2013). *Modelling covariates of infant and child mortality in Malawi* (Doctoral dissertation, University of Cape Town).

Mosley, W. H., & Chen, L. C. (1984). An analytical framework for the study of child survival in developing countries. *Population and Development Review*, *10*(0), 25-45.

Mugarura, A., & Kaberuka, W. (2015). Multilevel analysis of factors associated with child mortality in Uganda. *African Journal of Economic Review, 3(2), 125-139.*

Senayit, S., & Eshetu, W. (2013). Determinants of mortality among one to four years old children in Ethiopia: A study based on the 2011 EDHS data. *Ethiopian Journal of Health Development*, *27*(1), 8-15.

Seyoum, Y., & Sharma, M. K. (2014). Survival Analysis of Under Five Mortality in Rural Parts of Ethiopia. *International Journal of Statistics in Medical Research*, *3*(3), 266-281.

Solomon, G., & Emmanuel, G. (2016). Determinants of Under-Five Mortality in High Mortality Regions of Ethiopia: An Analysis of the 2011 Ethiopia Demographic and Health Survey Data. *International Journal of Population Research, 2016*.

Twum-Baah (1994) A study of infant, child and maternal mortality in Ghana. Ghana statistical services in collaboration with Ministry of Health and UNICEF, Accra, Ghana 24-32.

UNICEF. (2012). Levels and trends in child mortality. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation (IGME).

UNICEF. (2017). Levels and trends in child mortality. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation (IGME).

Van Raalte, A. A., Kunst, A. E., Lundberg, O., Leinsalu, M., Matikainen, P., Artnik, B., Deboosere, P., Stirbu, I., Wojtyniak, B. and Mackenbach, J. P. (2012). The contribution of educational inequalities to lifespan variation. *Population Health Metrics,* 2012, 10(3): 1-12.

Wang PD, Lin RS. Perinatal mortality in Taiwan. *Public Health* 1999; 113-33.

World Health Organization, & UNICEF. (2012). Trends in maternal mortality: 1990 to 2010: WHO, UNICEF, UNFPA and the World Bank estimates.

Yifang, H (2013). The analysis of child mortality in a developing country. Stockholm University. Unpublished master's thesis.

Appendix

Table A1: Summary of Influential Variables Statistics

Influential Variables	N	Minimum	Maximum	Mean	Std. Deviation
Analog of Cook's influence statistics	10641	.00000	.26988	.0032084	.01337195
Leverage value	10641	.00033	.19048	.0032891	.00540042
DFBETA for constant	10641	-.23530	.18304	-.0001015	.00653676
DFBETA for SEXC(1)	10641	-.00511	.00405	1E-7	.00087656
DFBETA for AGESM	10641	-.00025	.00022	0E-7	.00002612
DFBETA for BTYPE1(1)	10641	-.03071	.02521	0E-7	.00178120
DFBETA for BORDER1(1)	10641	-.31618	.28718	-.0000012	.00679587
DFBETA for BORDER1(2)	10641	-.01519	.01759	0E-7	.00151503
DFBETA for BORDER1(3)	10641	-.01162	.01199	0E-7	.00129632
DFBETA for PBIRTH1(1)	10641	-.28738	.31939	7E-7	.00661419
DFBETA for PBIRTH1(2)	10641	-.01911	.01070	-2E-7	.00142788
DFBETA for PBIRTH1(3)	10641	-.01771	.00840	-2E-7	.00136356
DFBETA for MAGEB1(1)	10641	-.02804	.23530	.0001035	.00372862
DFBETA for MAGEB1(2)	10641	-.02678	.23530	.0001035	.00368258
DFBETA for DBFEED(1)	10641	-.00967	.00760	-2E-7	.00111802
DFBETA for FSIZE1(1)	10641	-.02303	.02133	2E-7	.00178581
DFBETA for FSIZE1(2)	10641	-.01405	.01074	0E-7	.00128760
DFBETA for MEDU(1)	10641	-.15775	.05908	-9E-7	.00413172
DFBETA for MEDU(2)	10641	-.15633	.05715	-8E-7	.00407225
DFBETA for MEDU(3)	10641	-.15210	.05465	-6E-7	.00418654
DFBETA for WINDEX1(1)	10641	-.01936	.01958	1E-7	.00140974
DFBETA for WINDEX1(2)	10641	-.01368	.02541	0E-7	.00162552
DFBETA for REGION(1)	10641	-.04115	.02862	-3E-7	.00243662
DFBETA for REGION(2)	10641	-.04323	.02307	-7E-7	.00230968
DFBETA for REGION(3)	10641	-.04218	.03020	-4E-7	.00245383
DFBETA for REGION(4)	10641	-.04416	.02254	-4E-7	.00230857
DFBETA for REGION(5)	10641	-.04139	.02219	-6E-7	.00222205
DFBETA for REGION(6)	10641	-.04491	.02592	-7E-7	.00245542
DFBETA for REGION(7)	10641	-.04527	.02296	-7E-7	.00242356
DFBETA for REGION(8)	10641	-.04298	.03757	-9E-7	.00261663
DFBETA for REGION(9)	10641	-.04041	.03537	-8E-7	.00254380
DFBETA for REGION(10)	10641	-.06019	.09048	-6E-7	.00370017
DFBETA for PORES(1)	10641	-.02222	.03662	-4E-7	.00187440
DFBETA for TOILET1(1)	10641	-.02560	.03041	-2E-7	.00185623
DFBETA for TOILET1(2)	10641	-.01251	.01168	1E-7	.00115986
DFBETA for SDRINKW1(1)	10641	-.00863	.00700	0E-7	.00093767
DFBETA for POD1(1)	10641	-.01584	.01121	-2E-7	.00123207
Valid N (listwise)	10641				