ADDIS ABABA UNIVERSITY

OFFICE OF GRADUATE PROGRAMS

COLLEGE OF NATURAL SCIENCES

DEPARTMENT OF STATISTICS

COUNT REGRESSION MODELS FOR UNDER-FIVE DEATHS IN ETHIOPIA

YENEW ALEMU

A THESIS SUBMITED TO THE SCHOOL OF GRADUATE STUDIES OF ADDIS ABABA UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS

JUNE 2015

ADDIS ABABA UNIVERSITY

ADDIS ABABA, ETHIOPIA

Addis Ababa University

School of Graduate Studies

This is to certify that the thesis prepared by Yenew Alemu, entitled: **Count regression models for under-five deaths in Ethiopia** and submitted in partial fulfillment of the requirements for the Degree Master of Science in Statistics compiles with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the examining committee:

Examiner________________________________signature__________Date__________

Examiner________________________________signature__________Date__________

Advisor_________________________________signature__________Date__________

____

Chair of Department or Graduate Program coordinator

DECLARATION

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

Name: Yenew Alemu

Signature: ____________________

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

Date of submission: June, 2015

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

Mekonnen Tadesse (Ass. Prof.) ______________________

Advisor's Name Signature

Acknowledgment

It gives me a great pleasure to come to this final point to express my deepest gratitude to all individuals and institutions who contributed in one way or another to my study and the realizing of this thesis.

First of all I wish to express my profound gratitude to the Almighty God for his love, grace and protection throughout the course of my study in this university.

Secondly, Special thanks must go to my advisor, Mekonnen Tadesse(Ass.Prof.), for his invaluable advice and sincere guidance throughout my study. I thank him for a great number of things: for his inspiration, for his encouragement, for sharing his knowledge, for forcing me to dig deeper into my research and helping me to find the solutions. Without his sincere help, I think I would not be able to produce this quality thesis.

Thirdly, my deepest thanks go to all staff members of statistics department at Addis Ababa university especially, my instructors Dr. Dejen Tesfaw, Dr. Shibru Temesgen, Prof. Eshetu Wencheko, Prof. M. K. Sharma and Dr. Butte Gotu.

Finally I am indebted to all my families and classmates who were the source of strength towards the successful completion of the study.

Abstract

Count regression models for under-five deaths in Ethiopia

Yenew Alemu

Addis Ababa University, 2015

Under-five mortality is defined as the likelihood for a child born alive to die between birth and fifth birth day. Mortality of under the age of five has been the main target of public health policies and is a common indicator of mortality levels, especially in developing countries. It is also viewed as an indicator of the level of development, health and socioeconomic status of the population. The objective of this study was to identify determinants of under-five mortality in Ethiopia using the 2011 EDHS data. To achieve the objective of this study descriptive statistics and count regression models were used for data analysis using socio-economic, demographic and environmental related variables as explanatory variables and the number of under-five deaths per mother as the response variables. Factors influencing the number of under-five deaths have been identified. The study revealed that mother's age at the first birth, breastfeeding status, wealth index, current mother working, region and mother's level of education had statistically significant on the number of under-five deaths in rural parts of Ethiopia. Similarly, mother's level of education, age of mothers at the first birth, toilet facility and work/employment status of mothers were found to be statistically significant with the number of under-five deaths per mothers in urban parts of Ethiopia. Also region, age of mothers at the first birth, mother's level of education, breastfeeding status of mothers, wealth index and employment status of mothers were found to be statistically significant effect with the number of under-five deaths in Ethiopia.

Table of Contents

List of tables

List of Figures

LIST OF ABBREVIATIONS

CSA	Central Statistical Agency
EDHS	Ethiopia Demographic and Health Survey
U5M	Under-five Mortality
UNICEF	United Nations Children Fund
WHO	World Health Organization
BIC	Bayesian Information Criterion
AIC	Akaike Information Criterion
PRM	Poisson Regression Model
NBRM	Negative Binomial Regression Model
ZIP	Zero Inflated Poisson
ZINB	Zero Inflated Negative Binomial
NB	Negative Binomial
LRT	Likelihood Ratio Test
UN	United Nations
UNFPA	United Nations Population Fund
EHFP	Ethiopia Health and Family Planning
GTP	Growth and Transformation Plan

Chapter One
Introduction

1.1. Background of the study

Under-five mortality is defined as the likelihood for a child born alive to die between birth and fifth birth day. Mortality of under the age of five has been the main target of public health policies and is a common indicator of mortality levels, especially in developing countries (Gakusi and Garenne 2006). It is also viewed as an indicator of the level of development, health and socioeconomic status of the population. The study of under-five mortality has become one of the most important investigates of the developing countries including Ethiopia. There are two major reasons behind this: (i] high level of infant & child mortality and (ii] its relationship with fertility. The reduction of infant and child mortality indirectly helps in reducing fertility by decreasing the desired number of children to be born due to increased probability of survival of a child. Infant and Child mortality is a composite index reflecting environmental, social, economic, health care services and delivery situation on the one hand and maternal as well as family and community norms and practices on the other. Improving child survival has been a priority for both policy makers and health advocates worldwide.

Under-five mortality rate is defined as the probability (expressed as a rate per 1 000 live births) of a child born in a specified year dying before reaching the age of five years subject to current age-specific mortality rates (UNICEF 2008). It has several advantages as a barometer of child well-being in general and child health in particular. First, it measures an outcome of the development process rather than an input, such as per capita calorie availability or the number of doctors per 1,000 population – all of which are means to an end. Second, the U5MR is known to be the result of a wide variety of inputs: the nutritional status and the health knowledge of mothers; the level of immunization and oral rehydration therapy; the availability of maternal and child health services (including prenatal care); income and food availability in the family; the availability of safe drinking water and basic sanitation; and the overall safety of the child‟s environment, among other factors. Third, the U5MR is less susceptible to the fallacy of the average than, for example, per capita gross national income (GNI per capita). This is because the natural scale does not allow the children of the rich to be 1,000 times as likely to survive, even if the human-made scale does permit them to have 1,000 times as much income. In other words, it is much more difficult for a wealthy minority to affect a nation‟s U5MR, and it therefore presents a more accurate, if far from perfect, picture of the health status of the majority of children (and of society as a whole).

Infant and Child mortality is a widely accepted indicator of a country‟s progress in socio-economic and health development. Improvements in education, income, access to mass media, coverage and quality of health services, access to safe water and toilet facilities all play a role in reducing the infant and under-five mortality levels worldwide. In addition, under-five mortality is an important demographic, health and development issue for a number of reasons. It is a critical element in the calculation of overall mortality since the highest risk of death and proportion of deaths occur during childhood. It is one of the three measures (along with fertility and migration) that determine population size and growth rate, the age-sex distribution and the spatial spread of the population. Most importantly, of course, all human beings have a right to life and the extent to which that right is enjoyed is summarized by tracking trends and disparities in infant and child mortality.

Children are exposed to many hazards and risks as they grow and develop into adulthood, and unintentional injuries are the leading cause of death and disability for children and teenagers. The physical, social, cultural, political and economic environments in which they live can significantly increase or decrease their injury risks. The distribution of causes of childhood mortality for under-five children indicated that 34% of child deaths occurred in the first week after birth. The major causes of under-five mortality in Ethiopia were ARI/Pneumonia (21%), diarrhoea (14%), complications of prematurity (12%), intrapartum related events (birth asphyxia) (9%), meningitis (6%), and measles (4%). Other causes of death (including deaths due to severe malnutrition) accounted for 18% of under-five mortality (UNICEF, 2012).

Child deaths have been almost halved over the last few decades thanks to better nutrition, health care and standards of living. In 1990, more than 12 million children in developing countries died before the age of 5 from diseases such as diarrhea, malnutrition, pneumonia, AIDS, malaria, and tuberculosis. By 2012, that number had dropped to 6.6 million. Yet under-5 mortality rates remain unacceptably high, especially considering that most of these deaths are due to preventable or treatable causes. The World Bank is redoubling efforts in nutrition, health care, infrastructure, and other areas that can help save children‟s lives.

In Ethiopia, Demographic and Health Surveys (DHS) conducted since 2000 show decreasing trends in both infant and under-five mortality rates. Infant mortality was recorded as 97 deaths per 1000 live births in the year 2000 survey, 77 deaths per 1000 live births in the year 2005 survey and had declined to 59 per 1000 live births in the five year period leading up to the 2011 survey, a 39% reduction. Similarly, under-five mortality rate was measured as 166 per 1000 live births in the year 2000 survey, 123 deaths per 1000 live births in the year 2005 survey and reduced to 88 deaths per 1000 live births in the period leading up to 2011, a 47% reduction between 2000 and 2011 (EDHS 2000, 2005 and 2011 Final Reports).

By 2050, 37% of the world‟s children under age five will live in Sub-Saharan Africa; while close to 40% of all live births will take place in that region. Therefore the number of under-five deaths may stagnate or even increase without more progress in the region. Despite Sub-Saharan Africa‟s relatively high rates of under-five mortality, there are signs of progress in the region. The pace of decline in the under-five mortality rate has accelerated over time – increasing from 0.8 % per year in 1990–1995 to 4.2% per year in 2000–2013(WHO, 2014).

Mortality rates among children under the age of five remain strikingly high throughout the majority of sub-Saharan Africa. While other areas of the world have experienced declining rates of childhood mortality over the last 30 years, this area, for the most part, still maintains relatively high rates. It has been noted in 1995 that 18 of the 20 countries across the world with the highest childhood mortality rates were in sub-Saharan Africa (UN, 1995). As the world enters into the 21st century, childhood mortality remains a big issue for these developing countries, especially as researchers attempt to distinguish what factors contribute to the high levels.

Sub-Saharan Africa still has the world's highest rate of child mortality, despite reducing the number of under-five deaths by 48% since 1990. Sub-Saharan Africa is the region most affected and accounts for more than one-third of deaths of children under the age of five (Hill et al., 1999). In 2013, about half of global under-five deaths occurred in sub-Saharan Africa and 32 percent in South Asia. Around half of under-five deaths occur in just five countries: India (21%), Nigeria (13%), Pakistan (6%), Democratic Republic of the Congo (5%) and China (4%) (UN, 2014).

The situation in Ethiopian is also similar with that of the Sub Saharan Africa which is characterized by high child & infant mortality rate; it also ranks twenty seventh in the world by total number of death of infant. Almost one in every fourteen babies born in Ethiopia (69 per 1000 live births) doesn"t survive to celebrate the first birthday. Under five mortality in Ethiopia is also high (109 per 1000 live births) with one in every nine children dying before the fifth birthday (UNICEF, 2009). But data from the 2011 EDHS show that mortality has declined in Ethiopia over the past 15 years and that this decline is more pronounced over the last 10 years. There is a declining trend in infant and child mortality rate consistently over the past 15 years.

U5CM is one of the main problems in Ethiopia. In recent years, Ethiopia health and family planning (EHFP) has successfully implemented in a wide array of fertility and mortality reduction interventions. Besides, the growth and transformation plan (GTP) has been developed and under implementation starting from 2011 to improve access and quality of health services. However, despite all of these efforts, health care facilities in Ethiopia are limited and inadequate. Moreover, lack of health personnel, medicines and other facilities are not uniformly available. To expand our understanding about the most common and consistent factors on the risk of under-five child mortality, we have considered possible determinants of under-five child mortality using count regression model. Therefore, this study explores the Socio-demographic, environmental and socio-economic characteristics of under-five mortality in Ethiopia.

1.2. Statement of the problem

Despite the fact that a number of researchers have studied the risk factors of under-five mortality in Ethiopia, the majority of them studied the risk factors at national level. Such studies have overlooked an important point for policy makers as the findings at national level may not demonstrate the exact situation at the rural and urban areas separately. Under-five death is showing a declining trend over the last decades. Numbers of under-five death rates were 166 per 1000 live births in 2000 and 88 per 1000 live births in 2011. According to EDHS, 2011 report the level of under-five mortality in rural, urban and at national level are 114, 83 and 88 deaths per 1000 live births respectively. However, the rate is still very high and requires intervention to lower the death. The major causes of under-five deaths are early perinatal (death in first three days of life), neonatal tetanus, acute respiratory infection (ARI), malnutrition, diarrhoea, measles, and accidents in Ethiopia. Mostly in the rural, urban area separately and at the national level, there are limited studies that focused on the determinant factors of under-five mortality using count regression models. To understand the disparity in under-five mortality, we need to study the socio-economic, demographic and environmental factors.

Logistic regression has been widely used in analyzing child death data. However, since its dependent variable is dichotomized to be either “1” (death) or “0” (alive), logistic regression undercounts the total number of under-five mortality since multiple child deaths are collapsed into a single unit to fulfill the requirements of logistic regression. Logistic regression cannot provide sufficient information for studying the pattern of multiple child deaths. Therefore, we will develop and compare count regression models for the number of under-five deaths and discuss how they can enhance our understanding of the risk factors of under-five deaths.

1.3. Objectives

The general objective of the study is to identify socioeconomic, demographic and environmental factors of under-five mortality in Ethiopia using count regression model.

The specific objectives are

- To examine the determinants of under-five mortality in rural, urban parts of Ethiopia separately and at national level taking together using 2011 EDHS data.
- To fit an appropriate count regression model for rural, urban area separately and at national level taking together.

1.4. Significance of the study

- ✓ In order to give a better understanding of the factors affecting the under-five mortality in rural, urban area separately and at the national level.
- ✓ The results can provide an important input for any possible intervention in this area for the future.
- ✓ The study would serve as a guide to stakeholders in making informed and intelligent policy decisions with regard to under-five deaths and the management of the risk factors to avoid the death of children in the country.

1.5. Limitation of the study

The study has different limitations,

- The study used data from national surveys that have inherent gaps such as lack of data on children for women who had died though attempts were made to address them arising from the fact that only surviving women aged 15-49 years were interviewed.
- Some variables are not included because of large number of missing values like birth interval.
- The data used in this study are from the EDHS 2011. Thus, the results may not necessarily reflect the current situation of Ethiopia.
- Child specific biological variables are not considered in this study as we considered mothers as our response subject.

1.6. Definition of concepts

Mortality: the number of deaths in a given area or period or from a particular cause.

Under-five mortality (U5M): the probability of dying between birth and the fifth birthdays.

Infant-child mortality (IM): the probability of dying before the first birth day.

Neonatal mortality (NNM): the probability of a child dying between 0 and 28 days of age.

Post-neonatal mortality (PNN): the probability of dying after the first month of life but before the first birthday (the difference between infant and neonatal mortality).

Demographic characteristics: characteristics like age, marital status and breastfeeding status.

Socio-economic characteristics: characteristics like education, employment, residential area, wealth index and region.

Environmental characteristics: characteristics like source of drinking water and toilet facility.

Chapter Two
Literature review

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

Tibebu (2011) used data from 2005 EDHS to study risk factors and regional differentials in under-five mortality in Ethiopia by using multilevel count model. He found that mother education level, employment status of mother and economic status of mother were found to be statistically significant with under-five mortality. He also found that under-five child mortality differentials per mother among regions are significant.

Desta (2011) used data from 2000 and 2005 DHS to examine the socioeconomic, demographic and biological factors of infant and child mortality in Ethiopia using logistic regression analysis and found that marital status and mother"s educational levels are significant determinants of infant and child mortality. His finding showed that children born to mothers with more than 20 years of age at first birth are less likely to die before five years as compared to children born to mothers of age less than 15 at first birth. This study also suggested that infant and child mortality widely varied between regions in Ethiopia. However, most variation between regions does not have any significant effect on infant and child mortality.

Senayit (2012) used data from 2011 EDHS to study determinants of mortality among children aged one to five years using Cox proportional model and found that place of residence, mother"s age at first birth, mother"s education and marital status were found to be statistically significant with child mortality. The study showed that children living in rural areas face higher risk of mortality than children living in urban area. 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 is a significant determinant of child mortality showing that children born to illiterate mother experience higher risk of mortality than children born to mothers with primary and higher education. Similarly, the study by Bogaarts (2008) has shown that children belonging to educated mothers experience lower mortality in comparison with children belonging to uneducated mothers. In addition, the study by Van Raalte, Kunst, et al (2012) found that the mother‟s level of education on the survival of the child has been significant influence.

Amare et al (2007), using the data of structural questionnaire from Jimma town, Ethiopia, examined the determinants of under-five mortality of children in Gilgel Gibe field research Center, Southwest Ethiopia. They found that mother education and mother age were found to be statistically significant with under-five mortality. Under-five mortality was doubled in mother less than 20 years of age compared to above 20 years.

Ettarh and kimani (2011) used data from the 2008-2009 Demographic and Health Survey in Kenya to determine factors associated with mortality among under-five children in rural and urban areas by survival analysis. This study found that place of residence was a significant determinant of under-five mortality as seen in the hazard ratio of 3.6 for rural areas compared with urban areas. Mother‟s age was also a significant determinant of under-five mortality, with the highest likelihood of survival among children of mothers aged 20 years or more. Household wealth in Kenya was also a significant determinant of under-five mortality. In rural areas, households with greater wealth were less likely to have under-five deaths compared with the poorest households. This relationship was not seen among the urban households. In addition, the results of the study suggested that the probability of mortality within the first 5 years of life is significant effect to region of residence for rural children, but not for urban children.

Lemani (2013) used data from 2004 and 2010 DHS in Malawi to examine determinants of Infant and Child Mortality in Malawi using logistic regression and survival analysis. In this study, the 2010 data indicated that age of mothers and mother‟s education were found to be statistically significant with child mortality. Using 2010 DHS in survival analysis, infants from poor and middle income households were more likely to die as compared to those from rich households (OR=2.063 and 1.616, respectively).

Fitsum (2010) used data from 2005 EDHS to study socio-economic factors of early child hood mortality in Ethiopia. His finding showed that maternal education was a significant factor that affects childhood mortality negatively. The maternal age at first birth has significant and negative impact on childhood mortality.

Gebremariam (2001), studied diarrheal (which is major cause of morbidity and mortality of children in many developing countries) morbidity among young children in Eritrea using logistic regression model. He found that maternal education has a significant effect on child mortality.

Iram and Butt (2008) using the results of Pakistan Integrated Household Data (PIHD) to examine socioeconomic factors of childhood mortality in Pakistan. They found that the higher level of maternal education, the higher the infant and child mortality. They also found that maternal age at the first birth and working mother has significant and positive effect on childhood mortality.

Goro (2007) used data from 1993, 1998, and 2003 DHS in Ghana to examine determinants of infant and child mortality in three northern regions by using logistic regression model and found that only mother"s education had a significant impact for child mortality. Similarly, Yared (2013) used data from 2011 EDHS to examine the determinants of under–five mortality in rural parts of Ethiopia using survival analysis and found that mother"s educational level was found to be statistically significant with under-five mortality.

Based on Bangladesh demographic and health survey (2007) to examine the determinants of under-five mortality in Bangladesh using Chi-square test for independence and survival analysis, the study by Chowdhury (2013) has 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 the children born and brought up in urban areas.

Based on Kenyan Demographic and Health Survey Data (2003), Mustafa and Odimegwu (2008) examined the socioeconomic determinants of infant mortality in urban and rural Kenya using logistic regression model. The study has found that improving the socioeconomic status of the mother has the potential for reducing infant and under-five mortality rates. The predictors of the survival probability were dominated by biological and demographic factors (breastfeeding status) in both urban and rural areas. Similarly, Kyei (2011) found that the under-five mortality of South African children is affected significantly by socioeconomic factors such as the level of education of mother, marital status of the mother as well as her occupation and whether the mother‟s currently breastfeeding.

Mutunga (2004) used data from 2003 DHS in Kenya to investigate the determinants of socioeconomic and environmental variables on infant and child mortality in urban areas of Kenya. The results of his study indicated that living in wealthier households and having access to drinking water and sanitation facilities were weak predictors of infant and child mortality. However, mother‟s age and mother‟s education level had no significant effect on child mortality. Also, Hill (2001) in Kenya found that mother„s educational levels and economic status have a significant effect on infant and child mortality with urban areas was high risk of infant and child mortality than rural areas.

Based on the analysis of the 2003 SADHS, the study by Worku (2011) has found that improving the socioeconomic conditions of the mother in South Africa minimizes under-five mortality among South African children. The study has shown that the survival of under-five children in South Africa is significantly influenced by the breastfeeding status, the marital status of the mother, and toilet facility. The results also revealed that under-five mortality is higher among rural, uneducated and poor South African mothers.

Mutunga (2007), using data from 2003 KDHS employed survival analysis to examine the environmental determinants of child mortality in Kenya using survival analysis. The results from both the Weibull and Cox models have shown that households‟ socio-economic and environmental characteristics do have significant impact on child mortality. Of the demographic variables, the maternal age at first birth has significant and negative impact on childhood mortality. This means that the older the women are when giving their first birth, the higher the chances are for the children to survive their fifth birthday. Mothers those who drink protected water and mothers those who have toilet facility had lower mortality than those who drink unprotected water and those who have no toilet facility, respectively.

Mesike and Mojekwu(2012), using data from Demographic and Health Survey (DHS) 2008 and the annual abstract statistics of the National Bureau of Statistics (NBS) 2009 in Nigeria examined Environmental Determinants of Child Mortality by using principal component analysis and simultaneous multiple regression. As for the socio-economic variables, better survival prospect were found to exist in homes with high income. As expected, environmental characteristics of the household has been found to be significantly related to child mortality.

Using data from the Kenya 1989 population census, Gideon (2000) examines differences in under-five mortality using a count-data regression model. He found that mother‟s level of education, marital status, type of toilet facility, source of water and place of residence were found to be statistically significant in all the provinces of Kenya.

Klaauw and Wang (2003) developed a flexible parametric hazard rate framework for analyzing child mortality in rural Indian. Their model predicts significant effect between child mortality and access to provision of sanitation facilities and improving maternal education. Jacoby and Wang (2003) in a related study, examine the linkages between child mortality, morbidity, and household quality and community environment in rural China using a competing risks approach. Their findings among others showed that source of drinking water and maternal educations are significant predicator of child mortality. In examining the environmental determinants of child mortality in Ethiopia, Wang (2003) constructed three hazard models (the Weibull, the piecewise Weibull and the Cox model) to study three age-specific mortality rates by place of residence, mother education level, income status and access to basic environmental services (water and sanitation). Also, Espo (2002) in his study used indirect methods to estimate the levels and trends of mortality in Malawi. The results indicated that source of drinking water is a significant predictors of child mortality.

Bello and Joseph (2014) in their study examined the determinants of child mortality in Oyo State, Nigeria using logistic regression model. Their findings revealed that Breastfeeding status of mothers was a significant factor for child mortality in Atiba Local government Area of Oyo state while age of mother at first birth was not significant determinant.

Logistic regression model is applied by Sampson (2014), using the 2011 Ghana DHS to examine the determinants of under five mortality. His findings suggested that breastfeeding was the main determinants of under-five mortality in the Tano South district. On the other hand, other factors including mothers‟ education, mothers‟ occupation, household income and marital status did not show significant effect on under-five mortality.

Mondal et al. (2009) used the logistic regression model to investigated factors influencing infant and child mortality in Rajshahi District of Bangladesh. Findings revealed that the most significant predictors of neonatal, post-neonatal and child mortality levels are mother‟s age at birth. Similarly, Uddin et al. (2009) in their study investigated child mortality in Bangladesh also using the logistic regression. Results of analysis showed that occupation of mother, economic status of households and breastfeeding status were significant determinants of child mortality in Bangladesh.

Abimbola et al (2012) in their study examined the determinants of child mortality in rural Nigeria employing the 2008 Nigeria Demographic and Health Survey (NDHS) data. Data were analyzed using the Logit regression model. The result of analysis showed that education of mother and age of mother at first birth were among the significant factors influencing child mortality in rural Nigeria.

Gideon (2012) using the Poisson regression model, investigated under-five mortality differentials in urban East Africa: a study of three capital cities (Nairobi, Dar-es-Salaam and Kampala). The findings from these analyses underscore the fact that differences exist in under-five mortality levels in the three East African cities. Mothers age at the first birth, mother‟s level of education, marital status, and household living conditions (water and toilet facility) were significant determinants of infant and child mortality in all three cities.

Ermias (2013) used data from the demographic and health survey (DHS) conducted in Ethiopia 2005. The analysis was conducted using Cox proportional hazards model which analyses the effects of covariates on child mortality in rural Ethiopia. The study showed that source of drinking water, wealth index of household, mother‟s education have significant contribution on child mortality. In addition to these factors, age of the mother at first birth and quality of water are the factors that found to significantly affect childhood mortality (Mulugeta, 2012).

Pandey et al. (2012), employed Cox frailty model to investigate the determinants of under-five mortality in rural empowered action group states in India and the data from 2005-2006 National Family Health Survey (NFHS) was used. The result showed that mother‟s age at first birth, mother‟s education and mother‟s currently breastfeeding were significant determinants of under-five mortality. On the other hand, maternal occupation and household wealth index were not significant.

Logistic regression model is applied by Kouame (2014), using 2011-2012 Coted‟Ivoire‟‟ Demography Health Survey to investigate determinants of regional disparities in under age five mortality in Cote d'Ivoire. The results of analysis showed that maternal education, mother‟s age at the first birth and mother‟‟s marital status were significant effect on under-five deaths. This also study revealed a significant variation of under-age five mortality rate across region in Cote d‟Ivoire.

Antai (2011), from world health center, analyzed the Nigerian National Demographic and Health Survey (2003) data by using Multilevel Cox proportional hazards model to investigate regional inequalities in under-5 mortality in Nigeria. This study showed that under-5 mortality was found to be statistically significant with region of residence, with higher risks of under-5 deaths for children of mothers resident in the South region after adjusting only for individual-level risk factors and only community-level risk factors, as well as simultaneously adjusting for individual- and community-level risk factors.

Hammer et al (2006) argued that roughly 40% of under-five deaths occur during the neonatal period and this proportion was said to be considerably lower for regions with high absolute rates such as the sub-Saharan Africa. In Africa, infant and child mortality rates vary substantially from one sub-region to the other (Lutambi et al, 2010). The rates are as high as 163.2 and 197.6 deaths per 1000 live births in Guinea and Niger republic, respectively; and as low as 28.3 deaths per 1000 live birth in Egypt (Macro International Inc, 2011).

Using Ghana Demographic and Health Survey data of 1998 and World Bank data of 2000, Buor (2002) examined the effect of mothers‟ education on childhood mortality in Ghana by using logistic regression model and found that the higher level of mothers‟ education, the lower the infant and child mortality. In contrast, Anderson et al (2002) study among the African and Colored Population in South Africa indicates that, mother‟s education notwithstanding, environmental factors such as source of domestic water and type of sanitation significantly influence infant and child survival.

Frank et al (2001) used data from the 1994 Population and Housing Census of Ethiopia to identify the disk factors of infant and child mortality in urban Ethiopia by using logistic regression model. The results showed that the educational level and the well-being – in an economic and sanitary sense – of the household are the most important variables. Children of poor and illiterate women are the most disadvantaged groups in terms of infant and child mortality. This observation is more pronounced for women living in small towns. Other socio-demographic and socioeconomic characteristics are often directly related to the educational level of the household head or the educational level of mothers.

Rafiqul et al. (2013) based on data from Bangladesh demographic and health survey-2004 by used survival analysis and found that estimated under-five mortality is higher for not working mothers, rural areas and illiterate mothers than that of working mothers, urban areas and literate mothers respectively.

Tarihu and Eshetu(2013) used survival analysis to estimate social determinants of under-five mortality in Ethiopia The results showed that Being teenage mother at birth and child births to mothers residing in poor households were identified risk factors for increased under-five mortality in Ethiopia.

Reuben (2012) examined determinants of under-five mortality in Kenya during upsurge and declining period. The study findings have showed that in both KDHS 2003 and 2008, there was a significant variation on the under-5 mortality between children in households with improved sources of drinking water and those with a non-improved source of drinking water. Also, mother"s age at child birth significantly affects under-5 mortality in both datasets.

Using Ethiopia Demographic and Health Survey data of 2005, Bariagaber (2013) examined the housing determinants of under-five mortality in urban Ethiopia. The findings of this study showed that unprotected water and with access to toilet facility of the households are higher associated with under-five mortality.

Tettey (2003) in a study in Ghana and Nigeria observed that place of residence has a significant relationship with infant and child mortality in Ghana. The study showed that both infant and child mortality rates are higher in the rural areas; that is 67 deaths per 1000 live births than in the urban areas of Ghana – 43 deaths per 1000 live births.

Ezra and Gurum (2002) employed a logistic regression model to investigate Breastfeeding, birth intervals and child survival in southern Ethiopia. They revealed that children born to illiterate mother experience higher risk of mortality than children born to mothers with primary and above education.

Matthias et al (2012) used logistic regression analysis to examined socio-demographic determinants of mortality in hospitalized under-five children at a secondary health care centre in the Niger Delta. They found that significant risk factors associated with mortality were old maternal age, low maternal educational status and lower socioeconomic class. In addition, Buwembo (2010) using DHS 1997 and 2002 to investigate factors associated with under-five mortality in South Africa. This finding showed that mother"s education was significantly associated with under-five mortality.

Abdullah (2014) used data from DHS from Bangladesh to examine determinants of under-five child deaths per mother using zero-inflated regression models. He found that mother‟s level of education and wealth index are significant factors for under-five morality.

Based on the EDHS 2011 to determine statistically the correlates of child mortality in Ethiopia using logistic regression model, the study by Kasahun and Teshome (2014) found that maternal education, mother‟s age at the first birth and toilet facility are significant contribution on under-five child deaths.

Mohammed (2013), the data used from Cairo University Specialized Hospital and Benha University Hospital to examine supportive strategies regarding deaths Prevention for mothers of children under five years old using chi-square test. This finding showed that mothers age at the first birth, mother level of education and mother currently working were significant factors for under-five deaths.

Tesfaye (2011) used data from EDHS 2005 to investigate the determinants of under-five mortality per mother in Ethiopia using bayesian approach. He found that mother age at the first birth, mother level of education, current breastfeeding status of mother and source of drinking water were found to be statistically significant with under-five mortality. However, married women and mothers from rich and medium households were higher risk factors of under-five mortality than unmarried women and mothers from poor households, respectively.

Chowdhury et al. (2010) examine determinants of socio-economic determinants of neonatal, post neonatal, infant and child mortality per mother in Bangladesh using logistic regression analysis. They revealed that mothers level of education and occupation of mothers were statistically significant with neonatal, post neonatal, infant and child mortality.

Chapter Three
Data and methodology
3.1 Source of the data

The source of the data in this study is the 2011Ethiopia Demographic and Health Survey (EDHS). The 2011Ethiopia Demographic and Health survey was conducted by the Central Statistics Agency (CSA) with support from the ministry of Health. This is the third Demographic and Health Survey (DHS) conducted in Ethiopia, under the worldwide measure DHS project, a USAID-funded project providing support and technical assistance in the implementation of population and health surveys in countries worldwide.

The primary purpose of the EDHS is to provide policymakers and planners with detailed information on fertility, family planning, infant, child, adult and maternal mortality, maternal and child health, nutrition and knowledge of HIV/AIDS and other sexually transmitted infections along with other household characteristics in the nine regions and two administrative regions both at rural and urban levels. The survey interviewed a nationally representative population in about 18,500 households, and all women aged 15-49 and all men aged 15-59 in these households. The analysis presented in this study on under-five mortality was based on the 10,475 women aged 15-49 years.

During the analysis stage, Statistical Package for Social Science (SPSS) version 16, South Texas Art Therapy Association (STATA) version 12 and Microsoft-Excel were used as tools of analysis.

3.2 Variables in the study

The dependent variable for this study is the number of deaths of under- five deaths per mother. Based on the Mosley and Chen (1984) determinants of childhood morbidity and mortality framework for developing countries, experiences from the available similar studies and available data on the subject, the main predictors explored for under-five mortality have been grouped into demographic, socioeconomic and environmental factors.

3.3 Coding and Description of the variable

Detail description and coding of the demographic, socioeconomic, and environmental factors to under-five mortality is presented as follows.

Table 3.1: coding and description of demographic factors

No.	variable	levels	Descriptions of variables
1	MAGEB	0=<15 1=15-19 2=>=20	Mother‟s age at the firth birth
2	CBF	0=No 1=Yes	Mother‟s Currently Breastfeeding
3	MS	0=Currently married 1=Currently not married	Mother‟s Marital status

Table 3.2: coding and description of socio-economic factors

No.	variable	levels	Descriptions of variables
4	MEDU	0=No education 1=Primary 2=Secondary 3=Higher	Mother‟s level of education
5	CMW	0=No 1=Yes	Mother Currently working
6	REG	0=Tigray 1=Affar 2=Amhara 3=Oromiya 4=Somali 5=Benishangul-Gumuz 6=SNNP 7=Gambela 8=Harari 9=Addis Ababa 10=Dire Dawa	Region
7	WI	0=Poor 1=Medium 2=Rich	Household‟s wealth index
8	PRESD	0=Urban 1=Rural	Place of residence

Table 3.3: coding and description environmental factors

No.	variable	levels	Descriptions of variables
9	SDW	0=protected 1=unprotected	Source of drinking water
10	TF	0=has toilet facility 1=no toilet facility	Toilet facility

3.4 Methods of data analysis

3.4.1 Introduction

In this study, the variable of interest is a count variable. When the response or dependent variable is a count (which can take on non-negative integer values (0, 1, 2, ...), it is appropriate to use non- linear models based on non- normal distribution to describe the relationship between the response variable and a set of predictor variables. For count data, the standard framework for explaining the relationship between the outcome variable and a set of explanatory variables includes the Poisson and negative binomial regression models. Unlike linear regression, count data regression models have counts as the response variable that can take only nonnegative integer values.

Characteristics of count data:

1) Event counts are non-negative (lower bound is zero).
2) Counts are integers (discrete, rather than continuous variables).
3) A histogram will indicate a rapidly decreasing tail.
4) Distribution is not normal.

The two most popular models for count data are the Poisson model and the negative binomial model. The Poisson distribution should have the same mean and variance and the negative binomial regression model can be used instead of Poisson regression model when the data under consideration is over dispersed. A limitation of the Poisson distribution is the equality of its mean and variance. We may often observe count data processes where this equality is not reasonable: in particular, where the conditional variance is larger than the conditional mean. This is termed over-dispersion, and its presence renders the assumption of a Poisson distribution for the error process untenable.

However, when the data contain unobserved heterogeneity resulting from sources of over-dispersion, then the Poisson model will likely be inferior to the NB models. NB model allows the variance to differ from the mean. Over-dispersion is most often caused by highly skewed dependent variables. If the dispersion parameter (α) approaches to zero, it is appropriate to fit a Poisson regression model.

The negative binomial (NB) distribution is a two-parameter distribution. For positive integer n, it is the distribution of the number of failures that occur in a sequence of trials before n successes have occurred, where the probability of success in each trial is p. The distribution is defined for any positive n. The negative binomial distribution is a mixture of the Poisson distribution and the Gamma distribution. Unlike the Poisson, which is fully characterized by its mean μ, the NB distribution is a function of both μ and α. Its mean is still μ, but its conditional variance is $\mu(1 + \alpha\mu)$. As evident, as α closes to 0, the distribution becomes the Poisson distribution.

The negative binomial regression can be written as an extension of Poisson regression and it enables the model to have greater flexibility in modeling the relationship between the conditional variance and the conditional mean compared to the Poisson model. Also, an often encountered characteristic of count data is the number of zeros in the sample can exceed the number of zeros predicted by either Poisson or negative binomial model, and this is of interest because zero counts frequently have special status.

By over-dispersion, we mean that the variance of the outcome variable is larger than the expected value of the outcome variable. Zero-inflated means that there is excess number of zeros in the outcome variables. The ZINB model is useful for analysis of over-dispersed count data with an excess of zeros.

In practice, even after accounting for zero-inflation, the non-zero part of the count distribution is often over-dispersed. In this case, Greene .W.H (1994), described an extended version of the negative binomial model for excess zero count data, the zero-inflated negative binomial (ZINB) regression model, which may be more appropriate than the ZIP model. It has been established that the ZIP parameter estimates can be severely biased if the non-zero counts are over-dispersed in relation to the Poisson distribution. If count data are overdispersed such that the variance of the count variable is greater than the mean, then the Poisson assumption is violated. A negative binomial distribution may then be used for modeling purposes, as it uses an additional parameter in describing the variance of the count variable. If the data are still zero- inflated, a zero-inflated negative binomial (ZINB) model may be fit. Before attempting to develop and compare different count data models, it is very important to understand the assumptions made in the development of these models.

3.4.2 Poisson regression model

This regression model is a popular and simple regression model for count data. It assumes a Poisson distribution, characterized by a positive skewed and a variance equals the mean. Poisson regression analysis is a technique which allows to model dependent variables that describe count data (Cameron et al, 1998). According to Sturman (1999), the apparent simplicity of Poisson comes with two restrictive assumptions. First, the variance and mean of the count variable are assumed to be equal. The other restrictive assumption of Poisson models is that occurrences of the event are assumed to be independent of each other.

Poisson regression models provide a standard framework for the analysis of count data. Let Y_i represent counts of events occurring in a given time or exposure periods with rate μ_i. Y_i are Poisson random variables which the p.m.f. is characterized by

$$P(Y_i=y_i,\mu) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \quad , \mu_i>0,\ i=1,2,\dots n \ \text{ and } \ y_i = 0, 1, 2, \dots \qquad (3.1)$$

where, y_i denotes the value of an event count outcome variable occurring in a given time or exposure periods with mean parameter μ_i.

In Poisson model, the conditional variance is equal to conditional mean:

$$E(y_i)=Var(y_i)=\mu_i \tag{3.2}$$

This property of the Poisson distribution is known as equidispersion. Let x be an n×(p+1) covariate matrix. The relationship between Y_i and i^{th}row vector of x, x_i linked by $g(\mu_i)$ is

$$\ell n(\mu_i)=x_i^T\beta=\eta_i \tag{3.3}$$

This model is known as the Poisson regression or log-linear model. Where, $x_i = (1, x_{i1}, \dots, x_{ip})^T$ is the vector of explanatory variables and $\beta = (\beta_0, \beta_1, \dots, \beta_p)^T$ is the vector of the unknown regression parameters.

The regression parameters are estimated using the maximum likelihood estimation. The likelihood function of the Poisson model based on a sample of n independent observations is given by

$$\mathcal{L}(\beta, y_i) = \prod_{i=1}^{n} \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \tag{3.4}$$

The log-likelihood function is

$$\ell = \text{Log}(L(\beta)) = \sum_{i=1}^{n}[y_i \ln\mu_i - \mu_i - \ln y_i!] \tag{3.5}$$

The likelihood equation for estimating the parameter is obtained by taking the partial derivations of the log-likelihood function and setting them equal to zero. Thus, we obtain the first derivatives of ℓ with respect to the underlying parameters as follows:

$$\frac{\partial \ell(\beta)}{\partial \beta_j} = \sum_{i=1}^{n}(y_i - \mu_i)x_{ij} \tag{3.6}$$

If E(y_i) < var(y_i), then we speak about over- dispersion, and when E(y_i) > var(y_i), we say we have under-dispersion. Next, we will employ two tests of over dispersion where the Null Hypothesis (H_0) is: mean and variance of the response variable are equal against the Alternative Hypothesis (H_1): variance exceeds the mean. There are two basic criteria commonly used to check the presence of overdispersion :

1. Deviance, D(y, $\widehat{\mu}$), is given by

$$D(y, \hat{\mu}) = 2\times \sum_i^n \{y_i \ln(\frac{y_i}{\widehat{\mu}_i})\text{-}(y_i\text{-}\hat{\mu}_i)\} \tag{3.7}$$

where, y is the number of events, n is the number of observations and $\widehat{\mu}_i$ is the fitted Poisson mean.

2. Pearson chi-square test,x^2 is also given by

$$x^2 = \sum_{i=1}^n \frac{(y_i - \widehat{\mu}_i)^2}{\widehat{\mu}_i} \tag{3.8}$$

Over-dispersion may be a result of higher occurrence of zero counts and subject heterogeneity. If the model fits the data, both deviance and Pearson Chi-square statistics divided by the degrees of freedom are approximately equal to one. Values greater than one indicate the variance is an over-dispersion, while values smaller than one indicate an under-dispersion. It is possible to account for over-dispersion with respect to the Poisson model by introducing a scale (dispersion) parameter into the relationship between the variance and the mean (Pedan, 2001). Another way of checking the presence of over-dispersion is a statistical test of the hypothesis:

H0: α = 0 vs H_1: α > 0.

- If P-value of LRT_α < α (level of significant), which is an indicated of over-dispersion is present; negative binomial is preferred.
- If P-value of LRT_α > α (level of significant), Poisson regression is preferred.

The negative binomial regression model is more appropriate for over-dispersed data because it relaxes the constraints of equal mean and variance.

In the general Poisson regression model, we think of μ_i as the expected number of under five-child death from the i^{th} mother and the total number children ever born from the i^{th} mother is N_i. This means parameter will depend on the population size and the total number of children ever born from the individual mother. Thus the distribution of Y_i can be written as:

$$\mathbf{Y_i \sim poisson\ (N_i \mu_i)}$$

where $\mathbf{N_i}$ is the total fertility rate of i^{th} mother and $\boldsymbol{\mu_i} = \mathbf{exp}(\mathbf{X_i}^{T}\boldsymbol{\beta})$

The logarithm of the children ever born is introduced in the regression model as an offset variable. By including ιn[children ever born] as offset in the equation, it is differentiated from other coefficients in the regression model by being carried through as a constant and forced to have a coefficient of one (Gideon,2000). Thus, the GLM with an offset is given by

$$\boldsymbol{log\ \mu_i} = \boldsymbol{logN_i} + \boldsymbol{X_i}^{T}\boldsymbol{\beta}$$

The link between the expectation of the dependent variable and the linear predicator is a logarithmic function and the linear predictor contains a known part or offset. This allows for estimation of maximum likelihood, standard errors and the likelihood ratio goodness of fit chi-square statistics. The model suggests that both set of the parameters are dependent on the covariates. Furthermore, the number of children born will be equal to the observed deaths if the coefficients of the independent variables, denoted by $\boldsymbol{\beta}$, are all equal to zero. Since $\boldsymbol{logN_i}$ is a constant, any variation in the coefficients of the independent variables will show up affecting the dependent variable and not the number of children born. The procedure therefore allows us to obtain the maximum likelihood regression coefficients that can be easily interpreted in terms of differentials in the dependent variables. Using the negative binomial regression procedure, several regression equations are estimated to the relationship between under-five mortality changes when control variables earlier mentioned are introduced. Results from the negative binomial models are sometimes better expressed on more convenient scale.

3.4.3 Negative binomial regression model

This model is used when count data are overdispersed (i.e when the variance exceeds the mean). Overdisprsion, caused by heterogeneity or an excess number of zeros (or both) to some degree is inherent to most Poisson data. By introducing a random component into the conditional mean, the negative binomial regression model addresses the issue of over-dispersion. However, it equally models both zero and nonzero counts, which might result in a poor fit for data with excessive number of zeros. Therefore, it is always necessary to check the proportion of zero counts before developing a negative binomial regression model. We used the likelihood ratio test to determine the more appropriate model between the Poisson regression and negative binomial regression. Hilbe (2007) used negative binomial regression to model over dispersed Poisson data. When the negative binomial is used to model over-dispersed Poisson count data, the distribution can be thought of as an extension to the Poisson model. The negative binomial regression model uses a log link function between the dependent variable and independent variables. The only difference between the Poisson and the NB lies in their variances, regression coefficients tend to be similar across the two models, but standard errors can be very different. The NB regression model is

$$P(y_i; \mu i, \alpha,) = \frac{\Gamma(y_i + {}^1/_\alpha)}{y_i!\Gamma({}^1/_\alpha)}\left(1+\alpha\mu_i\right)^{\frac{-1}{\alpha}}\left(1+\frac{1}{\alpha\mu_i}\right)^{-y_i}, \quad y_i \geq 0 \quad \text{and} \quad \alpha > 0. \tag{3.9}$$

with mean and variance are given by

$$E(Y_i)=\mu_i=\exp(x_i{}^T\beta) \;\text{ and }\; Var(Y_i)=\mu_i(1+\alpha\mu_i) \tag{3.10}$$

where, α shows the level of overdispersion and $\Gamma(.)$ is the gamma function. If $\alpha = 0$,NB regression model will reduce to Poisson regression model. Often data will show over-dispersion (Variance>mean) or under-dispersion (Variance<mean). With over-dispersed data you could use the negative binomial regression model. This model adds unobserved heterogeneity by specifying

$$\mu_i=E(y_i)=\exp(x_i{}^T\beta)$$

Where, x_i^T is $1\times p$ row vector of covariate (including an intercepts), p is the number of covariate in the model and β is the corresponding $p\times 1$ column vector of unknown regression parameters. The likelihood function of the NB model based on a sample of n independent observations is given by

$$L(\mu, \alpha, y_i) = \prod_{i=1}^{n}\{\frac{\Gamma(y_i+{}^1/_\alpha)}{y_i!\Gamma({}^1/_\alpha)}\left(1+\alpha\mu_i\right)^{\frac{-1}{\alpha}}\left(1+\frac{1}{\alpha\mu_i}\right)^{-y_i}\} \quad (3.11)$$

The log-likelihood function ℓ of NB regression model is

$$\ell = \sum_{i=1}^{n}\{-\log(y_i!) + \sum_{k=1}^{y_i}\log(\alpha y_i - \alpha k + 1) - \left(y_i + {}^1/_\alpha\right)\log\left(1+\alpha\mu_i\right) - y_i\log(\mu_i)\} \quad (3.12)$$

where, $\frac{\Gamma(y_i+{}^1/_\alpha)}{y_i!\Gamma({}^1/_\alpha)} = \prod_{k=1}^{y_i}(y_i + \frac{1}{\alpha} - k) = \alpha^{-y_i}\prod_{k=1}^{y_i}(\alpha y_i - \alpha k + 1)$

The first-order derivatives of ℓ with respect to the parameters α and β are derived as follows:

$$\frac{\partial\ell}{\partial\beta} = \frac{\partial\ell}{\partial\mu}\frac{\partial\mu}{\partial\beta} = \sum_{i=1}^{n}\frac{(y_i-\mu_i)}{1+\alpha\mu_i}x_i \quad (3.13)$$

$$\frac{\partial\ell}{\partial\alpha} = \sum_{i=1}^{n}\{\sum_{k=1}^{y_i}\frac{y_i-k}{\alpha y_i-\alpha k+1} + \frac{\log(1+\alpha\mu_i)}{\alpha^2} - \frac{(y_i+{}^1/_\alpha)\mu_i}{1+\alpha\mu_i}\} \quad (3.14)$$

and the second derivatives are also given below:

$$\frac{\partial^2\ell}{\partial\beta\partial\beta^T} = -\sum_{i=1}^{n}\{\frac{(1+\alpha y_i)\mu_i}{(1+\alpha\mu_i)^2}\}x_i x_i^T \quad (3.15)$$

$$\frac{\partial^2\ell}{\partial\beta\partial\alpha} = -\sum_{i=1}^{n}\{\frac{(y_i-\mu_i)\mu_i}{(1+\alpha\mu_i)^2}\}x_i \quad (3.16)$$

$$\frac{\partial^2\ell}{\partial\alpha^2} = \sum_{i=1}^{n}\{(\sum_{k=1}^{y_i}\frac{-(y_i-k)^2}{(\alpha y_i-\alpha k+1)^2}) - \frac{2\log(1+\alpha\mu_i)}{\alpha^3} + \frac{2\mu_i}{\alpha^2(1+\alpha\mu_i)} + \frac{(y_i+{}^1/_\alpha)\mu_i^2)}{(1+\alpha\mu_i)^2}\} \quad (3.17)$$

For estimating regression coefficients β and dispersion parameter α, the Newton-Raphson iteration procedure is applied like Poisson model.

3.4.4 Zero- inflated model

In some cases, excess zeros exist in count data and considered as a result of overdispersion. In such a case, the NB model cannot be used to handle the over-dispersion which is due to the high amount of zeros. To do this, zero-inflation (ZI) can be alternatively used. These models assume that there are two latent classes of observations: those who can only have a 0 count (probability 1 for a zero), and those who have a positive probability for any count. In many applications this makes sense. Let y_i be a nonnegative integer-valued random variable and suppose $y_i = 0$ is observed with a frequency significantly higher than can be modeled by the usual model. Thus, the Zero inflated regression model is defined as:

$$P(y_i|xi) = \begin{cases} \omega_i + (1-\omega_i)f(0)\,, & y_i = 0 \\ (1-\omega_i)f(y_i)\,, & y_i = 1, 2, \dots, \quad 0 \le \omega_i \le 1 \end{cases} \tag{3.18}$$

Where, f(yi) follows either the Poisson or the Negative Binomial distribution.
The mean and variance of the zero inflated distribution $ZIf(y_i;\mu)$ distribution are given by

$$E_{zi}\,f(y_i;\omega,\mu) = (1-\omega)E_f(y_i;\mu)$$

and

$$Var_{zi}\,f(y_i;\omega,\theta) = (1-\omega)\left[E_f^2(y_i;\mu)\right] - [(1-\omega)E_f(y_i;\mu)]^2$$

$$= (1-\omega)\left\{Var_f(y_i;\mu) + \omega E_f^2(y_i;\mu)\right\}$$

Due to its complicated form, it is not easy to make comparisons between the variances of the underlying count data distribution and the zero-inflated models for the general case. Hence, we discuss each case separately, and we specify alternatives for the underlying count data distribution and the corresponding zero-inflated versions. From these alternatives the mixture with Poisson and the mixture with negative binomial are the most commonly used. Zero-inflated regression provides another way to handle excessive number of zeros. Zero-inflated regression also considers two data generating processes. However, instead of assuming all zero counts from a single generating process, zero-inflated regression assumes zero counts come from two different sources. Specifically, a zero count may come from the always-zero group or the not-always-zero group. Zero-inflated regression is also a two-part model. A logit model determines if a zero count is from the always-zero group or the not-always-zero group and a baseline model, either Poisson or negative binomial, governs both zero and positive counts from the not-always-zero group. In this study, we used the two zero inflated models. These are ZIP and ZINB regression models.

3.4.4.1 Zero- inflated Poisson regression model

The Zero-inflated Poisson regression study the relationship between dependent and independent variable(s) when there are many zeros value in the dependent variable, where the relationship is the mixture between Poisson model and Logistic model. Zero-inflated Poisson Regression also provides a flexible way of modeling zero counts and an attractive interpretation. The Zero-inflated Poisson, or ZIP, model is another model that one can use when the zeros in a dataset are argued to be caused by both chance and systematic factors (Min and Agresti, 2005). The transition stage addresses zero-inflation while the event stage addresses unobserved heterogeneity of responses including zeros (Jang, 2005). Welsh, Cunningham, Donnelly, and Lindenmayer (1996) refer to it as a mixture model. In order to explain the extra zeros in the variable y_i , The ZIP regression model is [Lambert, 1992],

$$P(y_i) = \begin{cases} \omega_i + (1-\omega_i)e^{-\mu_i}\,, & y_i = 0 \\ (1-\omega_i)\dfrac{e^{-\mu_i}\mu_i{}^{y_i}}{y_i!}\,, & y_i = 1,2,\ldots,\ \ 0 \le \omega_i \le 1 \end{cases} \tag{3.19}$$

where, Y_i~ZIP (μ_i, ω_i). The mean and variance of ZIP are given by

$$E(Y_i)=(1-\omega_i)\mu_i \ \text{ and } Var(Y_i)= E(Y_i)(\,(1+\omega_i\,\mu_i) \tag{3.20}$$

The parameters μ_i and ω_i can be obtained by using the link functions,

$$\log(\mu_i) = x_i^T\beta \quad \text{and} \quad \log(\frac{\omega_i}{1-\omega_i}) = z_i^T\gamma\,,\ i=1,2,\ldots..,n.$$

where, x_i^T and Z_i^T are covariate matrices, β and γ are the (p+1)× 1 and (q+1)× 1 unknown parameter vectors, respectively. The log-likelihood function of ZIP model is given by

$$\ell(\mu,\omega,y)=\sum_{i=1}^{n}\{\ln[\omega_i+(1-\omega_i)e^{-\mu_i}]I_{(y_i=0)}+[\ln(1-\omega_i)-\mu_i+y_i\ln(\mu_i)-\ln(y_i!)]I_{(y_{i>0})}\}$$

(3.21)

where, I (.) is the indicator function for the specified event, i.e. equal to 1 if the event is true and 0 otherwise.

To obtain the parameter estimates of ZIP regression models, $\widehat{\beta}$ and $\hat{\gamma}$, the Newton-Raphson method can be used. The first derivatives of with respect to β and γ are

$$\frac{\partial\ell}{\partial\beta_j}=\frac{\partial\ell}{\partial\mu_i}\frac{\partial\mu_i}{\partial\beta_j}=\sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{-(1-\omega_i)\mu_i e^{-\mu_i}}{\omega_i+(1-\omega_i)e^{-\mu_i}}\right]+I_{(y_{i>0})}(y_i-\mu_i)\}x_{ij} \quad ,\text{j=0,1,2,…,p;}$$

(3.22)

$$\frac{\partial\ell}{\partial\gamma_r}=\frac{\partial\ell}{\partial\omega_i}\frac{\partial\omega_i}{\partial\gamma_r}=\sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{(1-\omega_i)\mu_i e^{-\mu_i}}{\omega_i+(1-\omega_i)e^{-\mu_i}}\right]-I_{(y_{i>0})}[\frac{1}{1-\omega_i}]\}z_{ir} \quad ,\text{r=0,1,2,…..,q.}$$

(3.23)

The second derivatives of with respect to the underlying parameters as follows:

$$\frac{\partial^2\ell}{\partial\beta_j\beta_k}=\sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{-e^{-\mu_i}[(1-\mu_i)\omega_i+(1-\omega_i)e^{-\mu_i}](1-\omega_i)\mu_i}{[\omega_i+(1-\omega_i)e^{-\mu_i}]^2}\right]-I_{(y_{i>0})}\mu_i\}x_{ij}x_{ik},$$

j,k=0,1,…..,p; (3.24)

$$\frac{\partial^2\ell}{\partial\gamma_r\,\partial\gamma_s}=\sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{-(1-e^{-\mu_i})^2}{[\omega_i+(1-\omega_i)e^{-\mu_i}]^2}\right]-I_{(y_{i>0})}[\frac{1}{(1-\omega_i)^2}]\}z_{ir}\,z_{is} \quad ,\text{r,s=0,1,…,q.} \quad (3.25)$$

$$\frac{\partial^2\ell}{\partial\beta_j\gamma_r}=\sum_{i=1}^{n}\{I_{(y_i=0)}\left[\frac{e^{-\mu_i}\mu_i}{[\omega_i+(1-\omega_i)e^{-\mu_i}]^2}\right]\}x_{ij}z_{ir} \quad (3.26)$$

3.4.4.2 Zero-inflated negative binomial regression model

Zero-Inflated Negative Binomial (ZINB) regression is one of the methods used in troubleshooting overdispersion due to excessive zero values in the response variable (excess zeros). This model provides a way of modeling the excess number of zeros (with respect to a Poisson distribution or negative binomial distribution) in addition to allow for count data that are skewed and overdispresed. We used the vuong test, likelihood ratio based test, to compare the zero inflated negative binomial model with negative binomial regression model. A significant z-test indicates that the zero inflated models are preferred. The Zero-inflated negative binomial (ZINB) regression model assumes there are two distinct data generation processes. The result of a Bernoulli trial is used to determine which of the two processes is used. For observation *i*, with probability ω_i the only possible response of the first process is zero counts, and with probability of $(1-\omega_i)$ the response of the second process is governed by a negative binomial with mean μ_i. The zero counts are generated from the first and second processes, where a probability is estimated for whether zero counts are from the first or the second process. The overall probability of zero counts is the combined probability of zeros from the two processes. ZINB also arises in Bernoulli trials with non-equal success probabilities.

The overdispersed data are characterized by "excess zeros", "excess large outcomes" or both. ZINB model therefore accounts for "excess zeros" and also for extra heterogeneity in a positive outcome. The ZINB distribution is a general model for counts which nests the ZIP, NB, and Poisson models. We consider Y_i as a ZINB distribution. Specifically, we consider the distribution.

Gurmu and Trivedi (1996) used the zero-inflated negative binomial (ZINB) regression to model overdispersed data with an excess of zeros. This regression model was given by

$$P(y_i|\omega, \alpha, \mu) = \begin{cases} \omega_i + (1-\omega_i)(1+\alpha\mu_i)^{\frac{-1}{\alpha}}, & y_i = 0 \\ (1-\omega_i)\frac{\Gamma(y_i + 1/\alpha)}{y_i!\Gamma(1/\alpha)}(1+\alpha\mu_i)^{\frac{-1}{\alpha}}\left(1+\frac{1}{\alpha\mu_i}\right)^{-y_i}, & y_i > 0 \end{cases} \tag{3.27}$$

where, μ_i is the mean of the underlying negative binomial distribution, $\alpha > 0$ is the over dispersion parameter and is assumed not to depend on covariates and $0 \leq \omega_i \leq 1$. Also the parameters μ_i and ω_i depend on vectors of covariates x_i and z_i, respectively. The formulations for μ_i and ω_i are the same as those used in the zero-inflated Poisson regression model. In this case, the mean and variance of the Y_i are

$$E(Y_i) = (1 - \omega_i)\mu_i$$

$$Var(Y_i) = (1 - \omega_i)\mu_i(1 + \omega_i\mu_i + \alpha\mu_i)$$

ZINB approaches ZIP and NB as $\alpha = 0$ and $\omega_i = 0$, respectively. If both α and $\omega_i = 0$, then ZINB reduces to Poisson. The parameter μ_i is modeled as a function of a linear predictor, that is,

$$\mu_i = \exp(x_i^T\beta)$$

where, β is the (p+1)×1 vector of unknown parameters associated with the known covariate vector $x_i^T = (1, x_{i1}, \ldots, x_{ip})$, p is the number of covariates not including the intercept. The parameter ω_i, which is often referred as the zero-inflation factor, is the probability of zero counts from the binary process. For common choice and simplicity, ω_i is characterized in terms of a logistic regression model by writing as

$$\text{logit}(\omega_i) = \log\left(\frac{\omega_i}{1-\omega_i}\right) = Z_i^T\gamma$$

where, γ is the (q+1)×1 vector of zero-inflated coefficients to be estimated, associated with the known zero-inflation covariate vector $Z_i^T = (1, z_{i1}, \ldots, z_{iq})$, where q is the number of the covariates Z's not including the intercept. ZINB is also used to analyze exploratory data. When all the covariates are included in the log link model, as in the case of ZIP, the estimate of the inflated parameter was found to be zero.

3.5 Parameter Estimation of ZINB Model

The probability of the observed data, expressed as a function of the parameter is called the likelihood function. The maximum likelihood estimate of a parameter is the parameter value for which the probability of the observed data takes its greatest value. The ZINB distribution is not a standard GLM type exponential family distribution, even when the overdispersion parameter is known, and standard GLM fitting methods are not applied. To obtain the parameter estimates of ZINB regression models, $\hat{\alpha}$, $\widehat{\beta}$ and $\hat{\gamma}$, the Newton-Raphson method can be used. The log-likelihood function $\ell = \ell(\alpha,\mu_i,\omega_i;y)$, for the ZINB model is given below.

$$\ell = \ell(\alpha,\mu_i,\omega_i;y) = \sum_{i=0}^{n}\{I_{(y_i=0)}\log(\omega_i+(1-\omega_i)(1+\alpha\mu_i)^{\frac{-1}{\alpha}})+$$

$$I_{(y_i>0)}\log[(1-\omega_i)\frac{\Gamma(yi+\frac{1}{\alpha})}{y_i!\Gamma(\frac{1}{\alpha})}(1+\alpha\mu_i)^{\frac{-1}{\alpha}}(1+\frac{1}{\alpha\mu_i})^{-y_i}]\} \quad (3.28)$$

$$\text{since} \quad \frac{\Gamma(y_i+\frac{1}{\alpha})}{y_i!\Gamma(\frac{1}{\alpha})} = \prod_{k=1}^{yi}(y_i+\frac{1}{\alpha}-k) = \alpha^{-y_i}\prod_{k=1}^{y_i}(\alpha y_i-\alpha k+1)$$

Furthermore, ℓ can be written as

$$\ell = \sum_{i=0}^{n}\{I_{(y_i=0)}\log[\omega_i+(1-\omega_i)(1+\alpha\mu_i)^{\frac{-1}{\alpha}}]+ I_{(y_i>0)}[\log(1-\omega_i)-\log y_i!+\sum_{k=1}^{yi}(\alpha y_i-\alpha k+1) -$$

$$(y_i+\frac{1}{\alpha})\log(1+\alpha\mu_i)+\log y_i!+y_i\log\mu_i]\} \quad (3.29)$$

$$\frac{\partial\ell}{\partial\beta_j} = \frac{\partial\ell}{\partial\mu_i}\frac{\partial\mu_i}{\partial\beta_j} = \sum_{i=1}^{n}\{I(y_i=0)\frac{-(1-\omega_i)\mu_i(1+\alpha\mu_i)^{\frac{-1-\alpha}{\alpha}}}{\omega_i+(1-\omega_i)(1+\alpha\mu_i)^{-\frac{1}{\alpha}}} - I(y_i>0)(y_i+1/\alpha)\frac{\alpha\mu_i}{1+\alpha\mu_i}+y_i\}x_{ij} \quad (3.30)$$

$$\frac{\partial\ell}{\partial\gamma_r} = \frac{\partial\ell}{\partial\omega_i}\frac{\partial\omega_i}{\partial\gamma_r} = \sum_{i=1}^{n}\{I(y_i=0)\frac{[1-(1+\alpha\mu_i)^{-\frac{1}{\alpha}}]\omega_i}{\omega_i+(1-\omega_i)(1+\alpha\mu_i)^{-1/\alpha}} - I_{(yi>0)}\frac{\omega_i}{1-\omega_i}\}z_{ir} \quad (3.31)$$

$$\frac{\partial\ell}{\partial\alpha} = \sum_{i=1}^{n}\{I(y_i=0)\frac{\frac{-1}{\alpha}(1-\omega_i)\mu_i(1+\alpha\mu_i)^{\frac{-1-\alpha}{\alpha}}}{\omega_i+(1-\omega_i)(1+\alpha\omega_i)^{\frac{-1}{\alpha}}} + I(y_i>0)[\sum_{k=1}^{y_i}(\frac{y_i-k}{\alpha y_i-\alpha k+1}) + \frac{\log(1+\alpha\mu_i)}{\alpha^2}$$

$$-\frac{(y_i+\frac{1}{\alpha})\mu_i}{1+\alpha\mu_i}]\} \quad (3.32)$$

Newton-Raphson iteration procedure can be used for estimating the parameter of ZINB regression models.

3.6 Goodness of fit tests

3.6.1 Likelihood Ratio test

The Likelihood ratio test is a test of a null hypothesis H_0 against an alternative H_1 based on the ratio of two log-likelihood functions. The likelihood ratio test is a test of the overall model. The overall test statistic for likelihood ratio test is given as:

$$\text{Likelihood ratio test} = G^2 = -2(\ell_{null} - \ell_k) \sim x_{p-1}^2 \tag{3.33}$$

This statistic is called the likelihood-ratio test statistic.

Where: ℓ_{null} is the log-likelihood of the null model and ℓ_k is the log-likelihood of the model comprising k predictors, p is number of parameters and x_{p-1}^2 is a chi-square distribution with p-1 degree of freedom. If the test statistics exceeds the critical value, the null hypothesis is rejected. That means the overall model is significant.

In this study, to compare Poisson and NB regression models and also ZIP with ZINB regression models, we used significance of dispersion parameter and likelihood ratio (LR) test as criterions. The statistic of likelihood ratio test for α is given by the following equation:

$$LRT_\alpha = -2(LL_1 - LL_2) \tag{3.34}$$

This statistic has a Chi-squared distribution with 1 degrees of freedom and LL is log-likelihood. If the statistic is greater than the critical value then, the model 2 is better than the model 1.

3.6.2 Vuong Test

The Vuong test is a non-nested test that is based on a comparison of the predicted probabilities of two models that do not nest (Vuong, 1989). That means vuong test statistics are needed to provide the appropriateness of zero-inflated models against the standard count models. For instance, comparisons between zero-inflated count models with ordinary Poisson, or zero-inflated negative binomial against ordinary negative binomial model can be done using Vuong test. This test is used for model comparison. For testing the relevance of using zero-inflated models versus Poisson and NB regression models, the Vuong statistic is used. Let‟s define

$$m_i = \log\left(\frac{P_{1(Y_{i/X_i})}}{P_2(Y_{i/X_i})}\right)$$

where, $P_1(Y_{i/X_i})$ and $P_2(Y_{i/X_i})$ are probability mass functions of zero-inflated and Poisson or NB models, respectively. In general, $P_{N(Y_{i/X_i})}$ is the predicted probability of observed count for case i from model N, then the Vuong test statistic is simply the average log-likelihood ratio suitably normalized. The test statistic is

$$\mathrm{V} = \frac{\sqrt{n}\frac{\sum_{i=1}^{n} m_i}{n}}{\sqrt{\frac{\sum_{i=1}^{n}(m_i - \overline{m})^2}{n}}} = \frac{\sqrt{n}\ (\overline{m})}{\mathrm{s_m}} \qquad (3.35)$$

Where, $\overline{m}$ is mean of m_i, s_m standard deviation and *n* sample size.

The hypotheses of the Vuong test are:

$$H_o: E[m_i] = 0$$

$$H_1: E[m_i] \neq 0$$

The null hypothesis of the test is that the two models are equivalent. Vuong showed that asymptotically, V has a standard normal distribution. As Vuong notes, the test is directional (vuong, 1989).

- If V > $Z_{\alpha/2}$, the first model is preferred.
- If V < $-Z_{\alpha/2}$, the second model is preferred.
- If | V | < $Z_{\alpha/2}$, none of the models are preferred.

3.6.3 AIC and BIC

AIC and BIC are goodness of criteria used for model selection. The likelihood ratio test was used to compare the Poisson model and NB model. Many Monte-Carlo simulations indicate that the BIC and AIC selection criteria need to be used together [Dalrymple et al (2003) and Wang et al (1996)]. The model with smallest value of AIC or of BIC is preferable. Selecting an appropriate model can be used a standard likelihood information criteria, for example, Akaike information criteria (Akaike, 1973) or Baysians information criteria (Raftery, 1986) abbreviated by AIC and BIC, respectively, Where

$$\begin{aligned} \text{AIC} &= -2 \log \text{likelihood} + 2k \\ \text{BIC} &= -2 \log \text{likelihood} + k \ln(n) \end{aligned} \tag{3.36}$$

where, k = number of parameters and n = number of observations.

3.6.4 Test for individual predictors

Let β denote an arbitrary parameter. Consider a significance test of $H_0: \beta_0 = 0$. The simplest test statistic uses the large-sample normality of the ML estimator $\hat{\beta}$, let SE($\hat{\beta}$) denote the standard error of $\hat{\beta}$, evaluated by substituting the ML estimate for the unknown parameter in the expression for the true standard error. When H_0 is true, the test statistics

$$Z = \frac{\hat{\beta} - \beta_0}{\text{SE}(\hat{\beta})}$$

has approximately a standard normal distribution. Equivalently, z^2 has approximately a chi-squared distribution with df = 1. This type of statistic, which uses the standard error evaluated at the ML estimate, is called a Wald statistic.

The Wald statistic is

$$Z^2 = \left(\frac{\hat{\beta} - \beta_0}{\text{SE}(\hat{\beta})}\right)^2$$

Under H_0 true, Z^2 is a chi-square distribution with 1 degree of freedom. Wald statistics are for small samples. Likelihood-ratio tests are generally considered to be superior (Agresti, 2007).

Chapter Four

Results and Discussion

4.1 Descriptive statistics

Before proceeding to fit an appropriate count models, we make a descriptive analysis of the data in order to have an overall picture of the distribution of the number of under-five deaths. Thus we start with the description of the variables.

4.1.1 Number of under-five deaths per mother in rural parts of Ethiopia

Information on the number of deaths of under-five children obtained from a total of 8,668 women in the rural parts of Ethiopia was studied. Table 4.1 showed the frequency and percentage distribution of the number of under-5 deaths in rural parts of Ethiopia based on information from 8,668 women. While 64.5% of the women never experienced under-5 death of their children, 20.79%, 8.91%, 3.5%, 1.51%, 0.48%, 0.17%, 0.06% and 0.08% of them lost 1, 2, 3, 4, 5, 6, 7 and at least 8 of their under- five children, respectively.

Table 4.1: Frequency distribution of number of under-5 deaths in rural parts of Ethiopia

Number of deaths per mother	Frequency	percent
0	5591	64.50
1	1802	20.79
2	772	8.91
3	303	3.50
4	131	1.51
5	42	0.48
6	15	0.17
7	5	0.06
≥8	7	0.08
Total	8668	100.0

Summary statistics for the dependent variable used in the present study are presented in Table 4.2. The 8,668 observation values corresponding to each variable were used in the study. Among the 8,668 women, considered, 5,180 children died before the age of five (see Table A1). While the smallest value for number of under-five deaths was 0, the highest value was 15. Out of the 8,668 observed values used in the study, 5591 (64.5%) were zero (Table 4.1).

Table 4.2: Summary data for the dependent variable, number of under-5 deaths

	Minimum	Maximum	Mean	Variance	Skewness
Number of under-5 deaths	0.00	15.00	0.5976	1.052	2.537

Table 4.2 showed that the sample mean of the response variable, the number of under-five deaths was 0.5976 while the sample variance was 1.052. The fact that the mean is smaller than the variance, suggested a case of over-dispersion. Moreover, the data has excess zeros and thus one might expect that the Poisson model would not be appropriate to predict the number of under-five deaths. As shown in Figure 4.1, the distribution of the number of under-five deaths has a rapidly decreasing tail and is highly skewed to right with excess zeros.

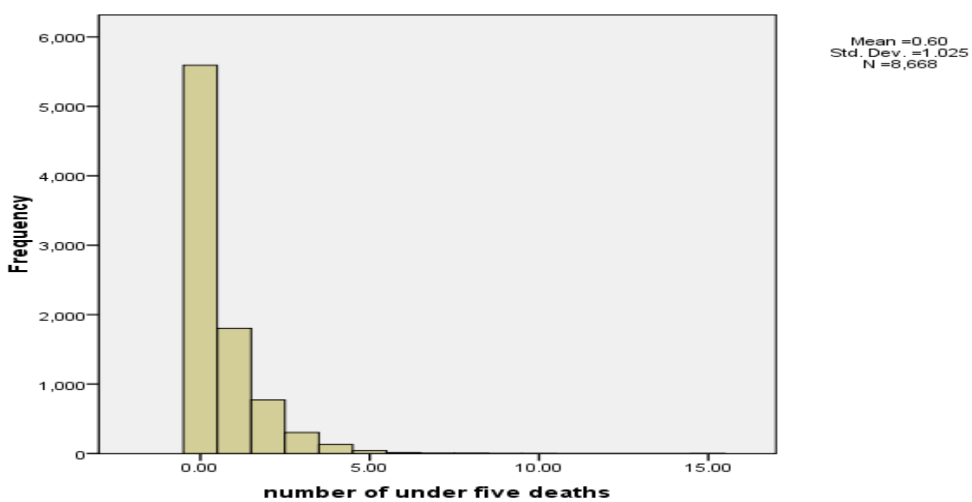

Figure 4.1: Histogram of the number of under-5 deaths

The descriptive results from the survey are presented in Table 4.3. The results indicate that rural resident mothers have higher mean number of under-five mortality than urban area. The women from Benshangul-gumuz have the highest mean of under-five child deaths when compared to other regions in rural parts of Ethiopia and at national level. But, the women from Somail have the highest mean of under-five mortality as compared to other regions in urban parts of Ethiopia. The youngest mothers (< 15 yrs) at the first birth have the highest mean number of under-five deaths and the mothers age at the first birth (>=20 yrs) have a lower mean number of U5M in rural and urban parts of Ethiopia and at national level.

Table 4.3: Descriptive results of the number of under-five child deaths per mother in rural, urban and at national level

variables	categories	rural		urban		National level	
		mean	S.D	mean	S.D	mean	S.D
Region	Tigray	0.4995	0.90994	0.3281	1.01259	0.4792	0.92385
	Affar	0.6125	1.09785	0.4576	0.85379	0.5945	1.07315
	Amhara	0.6204	0.9582	0.4302	0.87502	0.6061	0.95317
	Oromiya	0.578	0.97718	0.3056	0.75981	0.5532	0.96241
	Somali	0.5781	0.9581	0.6349	0.97866	0.5932	0.96339
	Benishangul-Gumuz	0.7805	1.35174	0.4474	0.94368	0.7522	1.32488
	SNNP	0.6286	1.01639	0.2152	0.69191	0.6064	1.0058
	Gambela	0.5512	1.01442	0.266	0.67482	0.5158	0.98289
	Harari	0.4582	0.86391	0.1019	0.42403	0.3361	0.76169
	Addis Ababa	-	-	0.0792	0.307	0.0792	0.307
	Dire Dawa	0.5714	0.93303	0.2187	0.63786	0.4306	0.84517
Mothers level of education	No education	0.6687	1.08021	0.5311	1.00728	0.6557	1.07423
	Primary	0.3775	0.79005	0.1755	0.50713	0.3252	0.73267
	Secondary	0.1515	0.47206	0.0483	0.23027	0.0674	0.29255
	Higher	0.069	0.25788	0.0878	0.32845	0.0847	0.31738
Source of drinking water	protected	0.5851	1.00691	0.2697	0.71833	0.4859	0.93731
	unprotected	0.607	1.03931	0.5426	0.96904	0.6058	1.03799
Type of toilet facility	Has toilet facility	0.5923	0.97861	0.2241	0.60517	0.4865	0.90298
	No toilet facility	0.6014	1.05821	0.5428	1.10423	0.5977	1.06116
Wealth index	Poor	0.6354	1.06897	0.6726	0.986	0.6362	1.06713
	Middle	0.5656	1.00002	0.1852	0.48334	0.5596	0.99497
	Rich	0.5298	0.92674	0.2591	0.71178	0.4067	0.84656
Mother's age at 1st birth	<15	1.1154	1.3597	0.7273	1.1657	1.0779	1.3464
	15-19	0.6108	1.01364	0.3487	0.82621	0.5744	0.9938
	>=20	0.4567	0.90717	0.1943	0.57636	0.3953	0.84871
Currently breastfeeding	No	0.6969	1.16344	0.2658	0.69049	0.5924	1.08393
	Yes	0.5581	0.96236	0.2979	0.76894	0.5214	0.94183
Current marital status	currently married	0.609	1.02192	0.3064	0.73519	0.5609	0.9881
	not currently married	0.5057	1.04976	0.1897	0.73112	0.4216	0.98484
Mother currently working	No	0.5794	0.98303	0.2444	0.63276	0.5307	0.94762
	yes	0.6456	1.12842	0.341	0.86019	0.5736	1.07875
Place of residence	urban	-	-	-	-	0.2839	0.73565
	rural	-	-	-	-	0.5976	1.02549

The results in Table 4.3 indicated that mothers with no education have the highest mean number of under-five child deaths when compared to mothers with primary, secondary and higher education in rural and urban areas separately and at national level taking together.

Table 4.3 showed that the breastfeeding mothers have a higher mean number of under-five child deaths as compared to non-breastfeeding mothers in rural parts of Ethiopia and at national level. But the reverse is true in urban area. In addition, working mothers have a higher number of under-five child deaths as compared to non-working mothers in rural and urban parts of Ethiopia separately and at national level. Moreover, married mothers have a higher mean number of under-five child deaths as compared to unmarried mothers in all study areas.

Among the women included in the study, mothers were from poor households have the highest mean number of under-five mortality as compared to other categories in rural and urban parts of Ethiopia separately and at national level.

With regards to type of toilet facility and source of drinking water, the higher mean number of under-five child deaths occurred, the mothers without protected water and toilet facility as compared to mothers used with protected water and toilet facility, respectively, in all study areas.

4.2 Comparison of Count Data Models in rural parts of Ethiopia

If the Deviance statistics and Pearson Chi-square statistic divided by their corresponding degrees of freedom are greater than one, then there is over-dispersion in the data suggesting Negative Binomial (NB) regression model.

Table.4.4: The results of over-dispersion test after fitting a Poisson regression

Statistics	Value	Degree of freedom	$\frac{\text{Value}}{\text{degree of freedom}}$	P-value
Deviance	9620.636	8648	1.11	0.000
Pearson chi-square	10027.21	8648	1.16	0.000

The test of over-dispersion, Deviance statistics and Pearson Chi-square Statistic divided by their corresponding degrees of freedom, are greater than one indicating over-dispersion. The likelihood-ratio test statistic is 348.32 with p<0.0001, we reject H_o that there is no over-dispersion, and conclude that there is significant over-dispersion in the data and the negative binomial regression model is favored over the regression model.

As shown in the summary Table 4.5, the likelihood-ratio chi-square values for all models were found to be significant. Thus, all regression models are significant. Next we compare all fitted models using log-likelihood, AIC and BIC values.

Table 4.5: Fit statistic of count regression models

criteria	Poisson	NB	ZIP	ZINB
Log-likelihood	-8412.6821	-8238.5233	-8253.059	-8227.931
-2LL	16,825.3648	16,477.0466	16,506.118	16,455.862
AIC	16865.36	16519.05	16548.12	16499.86
BIC	17006.71	16667.46	16696.53	16655.35
Likelihood ratio test	561.03 (0.000)	405.93 (0.000)	424.52 (0.000)	403.19 (0.000)

The model with the smallest AIC and BIC and the largest log likelihood is preferred. Since ZINB model has the smallest AIC and BIC and maximum log-likelihood, ZINB model is the most appropriate and preferred model among the four models. That is, the zero-inflated negative binomial regression model with the lowest value of AIC and the maximum value of log-likelihood is the most appropriate model for describing the number of under-five children deaths. Moreover, the fact that the LRT of α statistic, LRT_{α} = $-2(LLp-LL_{NB})$ =348.32, is greater than chi2(01), implies that the NB is better than the Poisson model and LRT_{α} = $-2(LL_{ZIP}-LL_{ZINB})$ = 50.25, greater than chi2(01),indicates that ZINB is better than ZIP model.

We used Vuong test to compare Zero-inflated regression models with other non-nested models, including Poisson regression and Negative Binomial regression models. The results indicated that Zero-inflated negative binomial (ZINB) regression model was the most appropriate count data model for this data. Although it is hard to distinguish Negative Binomial and Zero-inflated

Poisson (ZIP) regression models, they performed better than the Poisson regression model. The Vuong statistic values are presented in Table 4.6.

Table 4.6: Model comparisons by Vuong test for non-nested models

Model	Vuong Statistic(V)	Preferred model
ZIP VS Poisson	8.16	ZIP
ZINB VS NB	2.46	ZINB

Nested model: Models that are related where one model is an extension of the other.

Note:

If $V > 1.96$, the first model is preferred.

If $V < -1.96$, then the second one is preferred.

If $|V| < 1.96$, none of the models are preferred.

Table 4.7: Model comparisons by likelihood ratio test for nested models

Model	Likelihood ratio test (p-value)	Preferred model
NB VS Poisson	0.000	NB
ZINB VS ZIP	0.000	ZINB

Note:

H_0: The simpler model is preferred.

H_1: The more complex model is preferred.

If P-value<0.05, we reject H_0, and conclude H_1.

In addition, the following figure also confirms that the ZINB model is the most appropriate model among the four models considered.

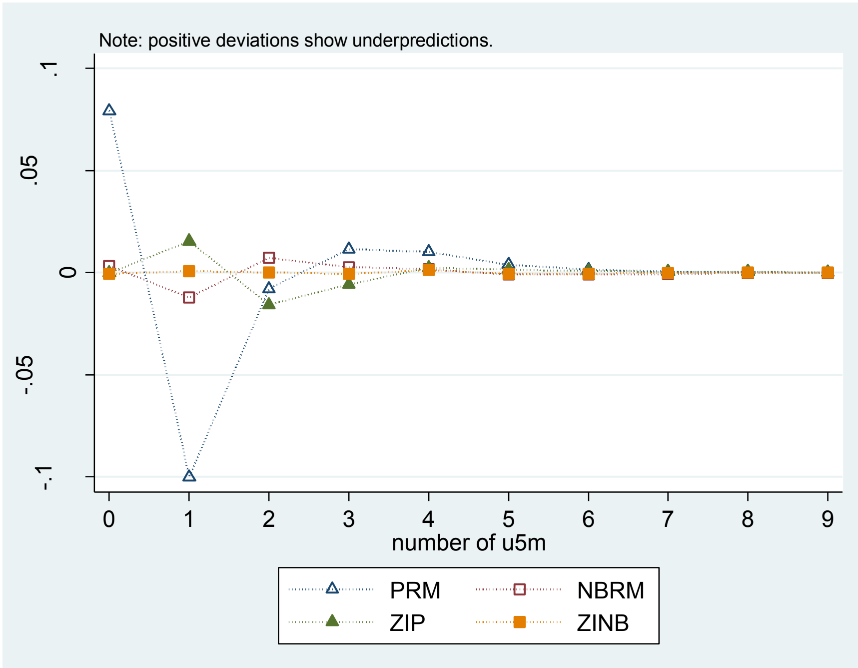

Figure 4.2 residual plots for estimated models

Moreover, the residual plot in Figure 4.2 showed that ZINB model fits well because almost all ZINB points pass through 0 and makes a straight line.

Table 4.8: Estimated coefficients of ZINB regression model

variables	Coef.	Std. Err.	z	P>|z|	[95% CI]
Region,Tigray(ref.)					
Affar	0.0416065	0.0724009	0.57	0.566	-.1002966 .1835095
Amhara	0.1411311	0.0690014	2.05	0.041	.0058908 .2763715
Oromiya	0.1569805	0.0657261	2.39	0.017	.0281598 .2858012
Somali	0.0403798	0.0779589	0.52	0.604	-.1124168 .1931763
Benshangul- Gumuz	0.3706548	0.0706173	5.25	0.000	.2322474 .5090622
SNNP	0.3081286	0.065627	4.7	0.000	.179502 .4367552
Gambela	0.2633426	0.0806256	3.27	0.001	.1053194 .4213659
Harari	-0.0685959	0.0986595	-0.7	0.487	-.261965 .1247732
Dire Dawa	0.1177358	0.092601	1.27	0.204	-.0637588 .2992305
Mother level of education, no education(ref.)					
primary	-0.4637706	0.0458781	-10.11	0.000	-.55369 -.3738512
secondary	-0.9691338	0.3312863	-2.93	0.003	-1.618443 -.3198247
higher	-1.6755	0.7211637	-2.32	0.02	-3.088955 -.2620451
Wealth index, poor(ref.)					
medium	-0.0869593	0.0434646	-2.00	0.045	-.1721482 -.0017703
rich	-0.0910793	0.0423884	-2.15	0.032	-.174159 -.0079995

Mother age at the first birth,<15(ref.)					
15-19	-0.4489802	0.0492739	-9.11	0.000	-.5455553 -.352405
20 and above	-0.6945656	0.0542082	-12.81	0.000	-.8008118 -.5883194
Currently breastfeeding, no(ref.)					
yes	-0.2534029	0.0347195	-7.3	0.000	-.3214518 -.185354
Current mother working, no(ref.)					
yes	0.0755776	0.0364916	2.07	0.038	.0040554 .1470999
_cons	-0.4335085	0.0812158	-5.34	0.000	-.5926885 -.2743285
log(children ever born)	1.000				
inflate					
Marital status, married(ref.)					
Unmarried	1.03762	0.2837085	3.66	0.000	.4815612 1.593678
_cons	-1.959543	0.3535019	-5.54	0.000	-2.652394 -1.266692
/lnalpha	-1.274886	0.2225431	-5.73	0.000	-1.711062 -.8387092
alpha	0.2794629	0.0621925			.1806738 .4322681

ref. = reference category of the variable.

4.3 Interpretation of ZINB regression model for positive counts in rural parts of Ethiopia

The results in Table 4.8 showed that region has a significant factor on the number of under- five deaths in the non-zero group. The expected number of under-five deaths for women from Amhara region was 1.15 times the expected number of under-five deaths for women in the reference group(Tigray) holding all other variables in the model constant. Similarly, the expected number of under-five deaths increased by a factor of 1.17 for women living in Oromiya compared to women in Tigray controlling for other variables in the model. In addition, the expected number of under-five mortality for women from Benshangul-gumuz, SNNP and Gambella had increased by a factor of 1.45, 1.36 and 1.33 as compared to the expected number of under five mortality in Tigray region, respectively, while holding all other variables in the model constant.

We can interpret mother‟s age at the first birth obtained in Table 4.8 using the reference category. The expected number of under-five deaths for women were decreased by a factor of 0.64 in the age group 15-19 compared to those in the age group less than 15 controlling for other variables in the model. Similarly, the expected number of under-five deaths for women decreased by a factor of 0.50 in the age group 20 and above as compared to those in the age group less than 15 controlling other variables in the model.

The finding of this study also revealed that mother‟s level of education had a significant factor on the number of under-five mortality. The expected number of under-five mortality for mothers with primary education was decreased by a factor of 0.64 as compared to those with no education (reference group) controlling other variables in the model. In addition, the expected number of the under-five mortality for mothers with secondary and higher level of education were decreased by a factor of 0.38 and 0.19 as compared to those with no education, respectively, controlling other variables in the model.

In this study, currently breastfeeding has a significant effect on the number of under-five deaths. The estimated number of under-five mortality for mothers who were breastfeeding is about 0.78 times lower than mothers who were not breastfeeding. In addition, mothers work status had a significant factor on the number of under- five deaths. The expected number of under- five deaths increased by a factor of 1.09 for working mothers as compared to that for non-working mothers while holding all other variables in the model constant.

According to the findings of this study, wealth index of the household has a significant influence on the number of under-five mortality. The expected number of under-five deaths for women in the medium and rich households was 0.92 and 0.91 times the expected number of under-five deaths for women in the poor households, respectively, while holding all other variables in the model constant.

4.4 Interpretation of ZINB regression model for covariates of zero counts in rural parts of Ethiopia

As shown in Table 4.8, mother‟s marital status has a significant effect on the probability of being an excess zero. The odds of being in the zero groups are increased by a factor of 2.82 for unmarried mothers as compared to married mothers controlling for other variables in the model.

4.5 Number of under-five deaths per mother in urban parts of Ethiopia

The distribution of socioeconomic, demographic and environmental background characteristics of mothers which affect under-five child mortality is presented in Table 4.9. Information on the number of deaths of under-five children obtained from a total of 1,807 women in the urban parts of Ethiopia was studied. From these, 328 women experienced 513 under-five deaths. Conversely, 81.8% of the women in urban areas never experienced under-5 death of their children. Nevertheless, 12%, 3.6%, 1.6%, 0.6%, 0.2%, 0.1% and 0.1% of them lost 1, 2, 3, 4, 5, 6 and 7 of their under–five deaths, respectively.

Table 4.9: Distribution of number of under-5 deaths in urban parts of Ethiopia

Number of deaths Per mother	Frequency	percent
0	1479	81.8
1	217	12.0
2	65	3.6
3	29	1.6
4	11	0.6
5	3	0.2
6	1	0.1
7	2	0.1
Total	1807	100.0

Descriptive statistics for the dependent variable used in the present study are presented in Table 4.10. The 1,807 observation values obtained for each variable were used in the study. Among the children of the 1,807 women considered in the study, 513 children died before the age of five. While the smallest value for the number of under-five deaths was 0, the highest value was 7. Out of the 1,807 observed values used in the study, 1,479 (81.8%) were zero.

Table 4.10: Summary data for the number of under-five deaths

	Minimum	Maximum	Mean	Variance	Skewness	kurtosis
Number of under-5 deaths	0.00	7.00	0.2839	0.541	3.658	17.723

As shown in Figure 4.3, the distribution of the number of under-five deaths has a rapidly decreasing tail and is highly skewed to right with excess zeros.

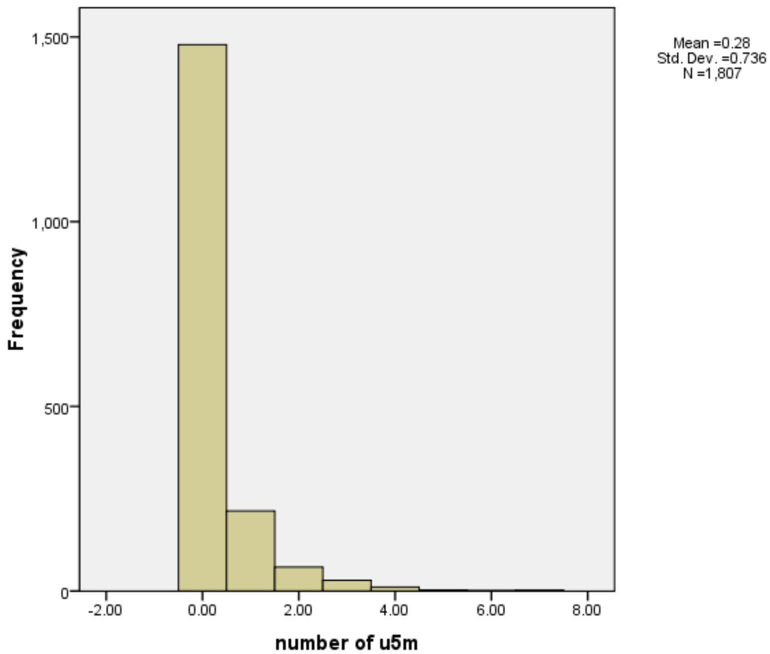

Figure 4.3: Histogram of the number of under-five deaths

4.6 Comparison of Count Data Models in urban parts of Ethiopia

For urban area, both the Deviance and Pearson statistics divided by the degree of freedom is approximately one in Table 4.11.

Table.4.11: Tests of over-dispersion

Statistics	value	Degree of freedom	$\frac{value}{degree\ of\ freedom}$
Deviance	1225.381	1789	0.70
Pearson Chi-Square	1870.536	1789	1.05

The formal test for significance of over-dispersion, the likelihood ratio test of α, $LRT_{\alpha} = -2 \times (LL_{Poisson} - LL_{NB})$, is computed. The likelihood ratio becomes 33.27, which corresponds to a p-value < 0.05, giving evidence of over-dispersion. Evidence of over-dispersion indicates inadequate fit of the Poisson model. In addition, we can identify more appropriate model by using AIC, BIC and log-likelihood.

Table.4.12: Model comparisons

Fit statistics	Poisson	NB
Log-likelihood	-986.83627	-970.20254
AIC	2009.673	1978.405
BIC	2108.662	2082.894

Table 4.12 showed that NB has maximum log-likelihood and minimum BIC and AIC indicating that the negative binomial model fits the data well.

Alternatively, we can identify the non-nested models by using vuong test. The value of Vuong statistic, V = 5.57 is greater than 1.96 implying that the ZIP regression model is better than the Poisson regression model. In addition, V = 4.54 is greater than 1.96 implying that the ZINB model is better than the NB regression model.

Table.4.13: Model comparisons by Vuong test for non-nested models

Model	Vuong statistics(V)	Preferred model
ZIP VS Poisson	5.57	ZIP
ZINB VS NB	4.54	ZINB

As shown in the Table 4.14, ZIP regression model has minimum AIC and BIC and maximum log-likelihood. Moreover, the fact that the likelihood ratio of α statistic, $LRT_{\alpha} = -2(LL_{ZIP} - LL_{ZINB}) = 1.41$, is less than chi2(01), implies that the ZIP is better than the ZINB regression model since the dispersion statistic($\alpha = 0.13$) is close to 0.

Table 4.14: Fit statistic for count regression models

Criteria	Poisson	NB	ZIP	ZINB
LL	-986.83627	-970.20254	-961.6857	-962.3918
AIC	2009.673	1978.405	1964.784	1965.371
BIC	2108.662	2082.894	2074.772	2080.859
LRT_{α} (p-value)	33.27(0.000)		1.41(0.1173)	

The ZIP regression model fit results are presented in Table 4.15.

Table 4.15: Estimated coefficients of ZIP regression model

variables	Coef.	Std. Err.	z	P>|z|	[95% CI]
Mother level of education, no education(ref.)					
primary	-0.6851459	0.118065	-5.8	0.000	-.916549 -.4537427
secondary	-1.644393	0.284829	-5.77	0.000	-2.202648 -1.086139
higher	-1.124482	0.3034474	-3.71	0.000	-1.719228 -.529736
Toilet facility, has toilet facility(ref.)					
No toilet facility	0.4553794	0.1026778	4.44	0.000	.2541347 .6566242
Mother age at the first birth,<15(ref.)					
15-19	-0.4387141	0.1579696	-2.78	0.005	-.7483289 -.1290994
20 and above	-0.6916586	0.1660779	-4.16	0.000	-1.017165 -.3661518
Current mother working, no(ref.)					
yes	0.4616567	0.0971435	4.75	0.000	.271259 .6520544
_cons	-0.6435852	0.1602475	-4.02	0.000	-.9576645 -.3295059
log(children ever born)	1.000				
inflate					
REG,Tigray(ref.)					
Affar	-1.249948	1.209543	-1.03	0.301	-3.620608 1.120713
Amhara	0.157807	0.68059	0.23	0.817	-1.176125 1.491739
Oromiya	0.0361875	0.6850303	0.05	0.958	-1.306447 1.378822
Somali	-1.367464	0.8548237	-1.6	0.11	-3.042888 .3079594
Benshangul- Gumuz	-0.1943995	0.7757449	-0.25	0.802	-1.714832 1.326032
SNNP	0.9807198	0.7509466	1.31	0.192	-.4911086 2.452548
Gambela	0.3106566	0.7214726	0.43	0.667	-1.103404 1.724717
Harari	1.768236	0.6238823	2.83	0.005	.5454497 2.991023
Addis Ababa	1.370265	0.585421	2.34	0.019	.2228606 2.517669
Dire Dawa	0.7982654	0.5829998	1.37	0.171	-.3443933 1.940924
Marital status, married(ref.)					
Unmarried	0.7145814	0.3060955	2.33	0.02	.1146451 1.314518
_cons	-0.7635159	0.5529071	-1.38	0.167	-1.847194 .3201622

4.7 Interpretation of zero inflated Poisson regression for covariates of non-zero groups in urban parts of Ethiopia

Mother‟s education level was an important factor for predicting the number of under-five mortality. The expected number of under-five deaths decreased by a factor of 0.50 for mothers with primary education compared to those with no education controlling other variables in the model. Also the expected number of under-five mortality for mothers with secondary and higher education was decreased by a factor of 0.19 and 0.32 as compared to those with no education, respectively, controlling other variables in the model.

As shown in Table 4.15, toilet facility had a significant effect on the number of under-five deaths. The expected number of under-five deaths increased by a factor 1.58 for a woman with no access to toilet facility as compared to a woman with access to toilet facility holding all other variables in the model constant.

Table 4.15 revealed that mother‟s age at the first birth has a significant effect on the number of under-five deaths. The expected number of under-five deaths for mothers in the age group 15-19 had decreased by a factor of 0.64 as compared to the expected number of under-five deaths for mothers in the age group less than 15 controlling for other variables in the model. In addition, the expected number of under-five mortality for mothers in the age group 20 and above was 50% less than that for mothers under 15 years of age controlling for other variables in the model.

Table 4.15 showed that mothers work status had a significant effect on the number of under- five deaths. The expected number of under- five deaths increased by a factor of 1.59 for working mothers as compared to that for non-working mothers while holding all other variables in the model constant.

4.8 Interpretation of ZIP regression for covariate of zero counts in urban parts of Ethiopia

Table 4.15 also shows that region has a significant effect on the odds of being in the always zero group. The odds of being in the zero group among children born to mothers in Harari and Addis

Ababa increased by a factor of 5.86 and 3.94 to the odds among children born in Tigray, respectively, holding all other variables in the model constant.

As shown in Table 4.15, mother‟s marital status has a significant effect on the probability of being an excess zero. The odds of being in the zero groups are increased by a factor of 2.04 for unmarried mothers as compared to married mothers controlling for other variables in the model.

4.9 Number of under-five deaths per mother in Ethiopia

The data to be analyzed for this study were obtained from Ethiopian Demographic and Health survey (EDHS) 2011. The initial population consisted of 11654 women under-fifty years of age. Of these, 10,475 women were selected and studied and the remaining were excluded due to incompleteness and inconsistency of data on the variables which are considered as important and relevant for the analysis. Among the 10,475 women, 67.5% never experienced under-five deaths of their children while 32.5% of the women experienced death of their under-5 children.

Table 4.16: Distribution of number under-5 deaths in Ethiopia

Number of deaths per mother	frequency	percent
0	7070	67.49
1	2019	19.27
2	837	7.99
3	332	3.17
4	142	1.36
5	45	0.43
6	16	0.15
7	7	0.07
<=8	7	0.07
Total	10,475	100.00

As shown in Table 4.16, 19.27%, 7.99%, 3.17%, 1.36%, 0.43%, 0.15%, 0.07% and 0.07% of the women lost 1, 2, 3, 4, 5, 6, 7 and at least 8 of their under-5 children, respectively. The following histogram showed that the frequency of the dependent variable is rapidly decreasing and highly skewed to the right with excess zeros.

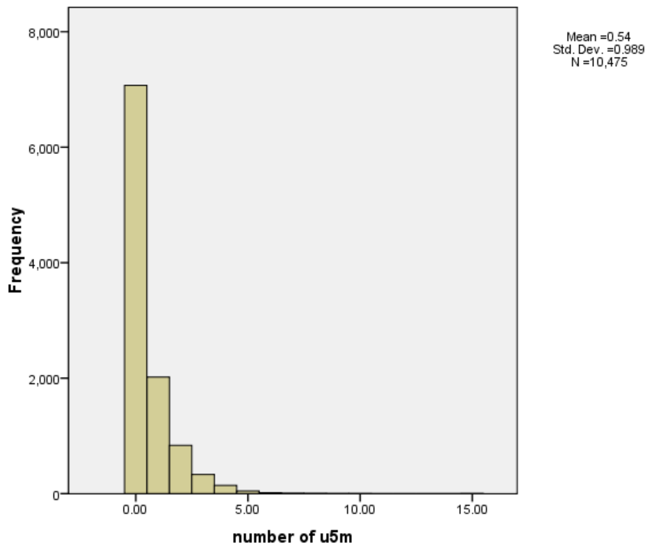

Figure 4.4: Histogram of the number of under-five deaths

Since 7070 (67.5%) of the data are zeros, the Zero-inflated regression model may be appropriate for analyzing this data set. As shown in Table 4.17, the variance of the outcome variable is greater than the mean and the ratio $\frac{0.978}{0.5435}$=1.799 > 1 indicating over-dispersion.

Table 4.17: Summary data for the dependent variable, number of under-5 deaths

	Minimum	Maximum	Mean	Variance	skewness
Number of under-five mortality	0.00	15.00	0.5435	0.978	2.67

4.10 Comparison of Count Data Models in Ethiopia

Both test statistics (Deviance and Pearson Chi-square) indicating over-dispersion. As shown in Table 4.18.

Table.4.18: The results of over-dispersion test

Statistics	Value	Degree of freedom	$\frac{\text{Value}}{\text{degree of freedom}}$	P-value
Deviance	10945.43	10453	1.05	0.0004
Pearson chi-square	11902.45	10453	1.14	0.0000

In addition, the likelihood-ratio test of dispersion statistic is 405.67 with $p<0.0001$ and we reject H_o that there is no over-dispersion, and concluded that there is significant over-dispersion in the data and the negative binomial is favored over the Poisson regression model.

The analysis of our data implied excess zeros (approximately 67.5% of the women never experienced under-five deaths), we fit Zero-inflated regression models in order to handle excessive number of zeros. Zero-Inflated count models provide a way of modeling excess zeros as well as allowing for over-dispersion.

In order to check excess number of zeros in this data set, we will use vuong test. The Vuong statistic has been computed to compare the PR model with the ZIP and the NB model with the ZINB. The Vuong statistics of ZIP against PR was obtained as 9.01 and determined as being significant ($p < 0.05$). The result indicates that Zero-inflated Poisson (ZIP) regression model is a more appropriate count data model compared to the Poisson model. Similarly, The Vuong statistics of ZINB against NB was obtained as 3.15 and is significant ($p < 0.05$). The result indicates that ZINB regression model is a more appropriate count data model compared to NB model.

As shown in the Table 4.19, the likelihood-ratio chi-square values for all models were found to be significant. Thus, all regression models are significant. But we identify the most appropriate model by using the following goodness of fit test. The model with maximum log-likelihood and smallest AIC and BIC was ZINB. Therefore, ZINB model is the most appropriate model.

Table 4.19: Fit statistic of count regression models

criteria	Poisson	NB	ZIP	ZINB
Log-likelihood	-9,449.2245	-9,246.3892	-9,259.633	-9,230.952
-2LL	18,898.449	18,492.7784	18,519.266	18,461.904
AIC	18,942.45	18,538.78	18,565.27	18,509.9
BIC	19,102.1	18,705.68	18,732.17	18,684.07
Likelihood ratio test	978.62 (0.000)	728.92 (0.000)	596.43 (0.000)	570.33 (0.000)

Table 4.20: Model comparisons by Vuong test for non-nested models

Model	Vuong Statistic(V)	Preferred model
ZIP VS Poisson	9.01	ZIP
ZINB VS NB	3.15	ZINB

In addition, the likelihood ratio test of α calculated in order to compare the models of NB vs Poisson and ZIP vs ZINB turned out to be very important ($p < 0.05$). LRT_{α} statistic =-2(LLpoisson-LL_{NB}) = 405.67, more than chi2(01),which indicates that NB is better than Poisson model. Similarly, we compared ZIP with ZINB by using likelihood ratio test. LRT_{α}=-2(LL_{ZIP}-LL_{ZINB}) = 57.36, more than chi2(01) thus, ZINB is more appropriate than ZIP model. Moreover, the following residual plot also confirms that ZINB is the most appropriate model among the four models considered because almost all the ZINB points pass at 0 and make a straight line.

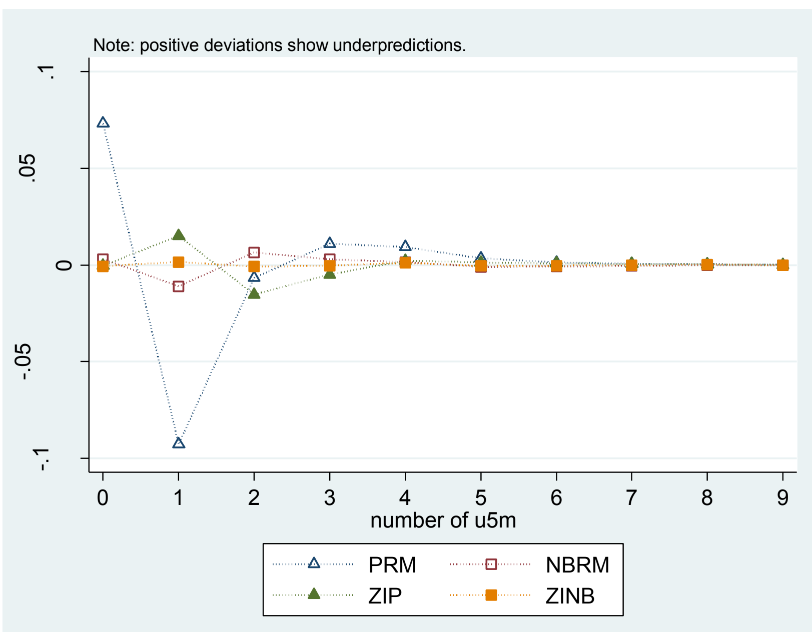

Figure 4.5: Residual plot for estimated model

The ZINB model fit results are presented in Table 4.21.

Table 4.21: Estimated coefficients of ZINB regression model at national level

variables	Coef.	Std. Err.	z	P>|z|	[95% CI]
Region, Tigray(ref.)					
Affar	0.0496609	0.0693262	0.72	0.474	-.0862159 .1855377
Amhara	0.1169031	0.0668408	1.75	0.08	-.0141024 .2479086
Oromiya	0.1262664	0.0634615	1.99	0.047	.0018843 .2506486
Somali	0.0925894	0.0707068	1.31	0.19	-.0459934 .2311722
Benshangul- Gumuz	0.3330045	0.0682984	4.88	0.000	.1991421 .466867
SNNP	0.2680209	0.0636909	4.21	0.000	.143189 .3928528
Gambela	0.2210483	0.0776142	2.85	0.004	.0689272 .3731694
Harari	-0.176413	0.0933599	-1.89	0.059	-.359395 .006569
Addis Ababa	-0.8193613	0.2037381	-4.02	0.000	-1.218681 -.4200419
Dire Dawa	0.0244292	0.0852551	0.29	0.774	-.1426678 .1915262
Mother level of education, no education(ref.)					
primary	-0.5016564	0.0430774	-11.65	0.000	-.5860865 -.4172263
secondary	-1.441526	0.2147657	-6.71	0.000	-1.862459 -1.020593
higher	-1.260998	0.2714506	-4.65	0.000	-1.793032 -.7289648
Wealth index, poor(ref.)					
medium	-0.0743468	0.0432806	-1.95	0.015	-.1691753 .0004816
rich	-0.0830559	0.0395122	-2.1	0.036	-.1604985 -.0056134
Mother age at the first birth,<15(ref.)					
15-19	-0.4499776	0.0477169	-9.43	0.000	-.543501 -.3564542
20 and above	-0.6971563	0.0521391	-13.37	0.000	-.7993471 -.5949654
Currently breastfeeding, no(ref.)					
yes	-0.2237582	0.0330238	-6.78	0.000	-.2884837 -.1590327
Current mother working, no(ref.)					
yes	0.1217118	0.0343011	3.55	0.000	.0544828 .1889408
_cons	-0.4429347	0.0791439	-5.6	0.000	-.5980538 -.2878156
log(children ever born)	1.000				
inflate					
Place of residence, urban(ref.)					
rural	-0.7909253	0.2600484	-3.04	0.002	-1.300611 -.2812398
Marital, married(ref.)					
Unmarried	1.07402	0.257775	4.17	0.000	.5687899 1.579249
_cons	-1.225963	0.294587	-4.16	0.000	-1.803343 -.6485831
/lnalpha	-1.224465	0.2091214	-5.86	0.000	-1.634336 -.8145948
alpha	0.2939148	0.0614639			.1950819 .4428187

4.11 Interpretation of ZINB model for covariates of positive counts in Ethiopia

Regarding the regional differences in under-five mortality, the expected number of under-five deaths for women from Oromiya, Benshangul-Gumuz, SNNP and Gambela were 1.13, 1.40, 1.31 and 1.25 times higher than the expected number of under-five deaths for women in Tigray, respectively, controlling for the other variables in the model. Conversely, the expected number of under-five deaths for women from Addis Ababa was about 0.44 times lower than that of Tigray controlling other variables in the model.

The finding of this study also revealed that mother‟s level of education had a significant factor on the number of under-five deaths. The expected number of under-five deaths for mothers with primary, secondary and higher education was about 0.61, 0.24 and 0.28 times less than that of the expected number of under-five deaths for mothers with no education, respectively, controlling for the other variables in the model.

When we look at the age of mothers at first birth, the expected number of under-five deaths for mothers aged 15-19 decreased by 36% as compared to the expected number of under-five deaths for mothers aged less than 15 (reference group) controlling other variables in the model. Also, the expected number of under-five deaths for mothers aged 20 and above decreased by 50% as compared to the expected number of under-five deaths for mothers aged less than 15 controlling other variables in the model.

Table 4.21 showed that mothers currently breastfeeding and currently working had statistically significant on the number of under-five deaths. The expected number of under-five deaths for breastfeeding mother was 20% less than the expected number of under-five deaths for non-breastfeeding mother controlling other variables in the model. Moreover, the expected number of under-five deaths for working mothers increased by 13% as compared to those non-working mothers while holding all other variables in the model constant.

According to the findings of this study, wealth index of the household has a significant influence on the number of under-five mortality. The expected number of under-five deaths for women in the medium and rich households was decreased be a factor of 0.93 and 0.92 times the expected number of under-five deaths for women in the poor households, respectively, while holding all other variables in the model constant.

4.12 Interpretation of ZINB model for covariates of zero counts in Ethiopia

As shown in Table 4.21, place of residence has a significant effect on the probability of being in the always zero group. The odds of being in the always zero group decreased by 55% for rural women as compared to those women in urban centers controlling other variables in the model.

Table 4.21 revealed that marital status have a significant factor on the probability of being an excess zero deaths. The odds of being in the always zero group increased by a factor of 2.9 for unmarried women as compared to married holding all other variables constant.

4.13 Discussion of the Results

This study was carried out to identify the risk factors of under-five mortality in rural and urban areas separately as well as Ethiopia as a whole based on EDHS 2011 data. The total number of women from rural areas included in the present study was 8,668 among which 35.5% experienced under-five deaths due to different factors. The most appropriate model was selected from four possible count models and the ZINB regression model was selected as the most appropriate model in rural parts of Ethiopia.

As the results show, the number of under-five mortality in rural Ethiopia has a wide variation among regions. This result is consistent with the findings by Ettarh and Kimani (2011).

In this study, mothers" education was found to be an important socio-economic predictor of the number of under-five mortality in rural Ethiopia. Under-five mortality decreased with increased level of mother"s education. This result is consistent with the findings by Yared (2013), Ermias (2013), Abimbola et al (2012), Klaauw and Wang (2003), Jacoby and Wang (2003) and Pandey et al. (2012).

The current study revealed that those children whose mother‟s age at first birth was below 15 years had the higher risk of dying relative to children whose mothers‟ age was 15 and above. A similar study in Kenya by Ettarh and Kimani (2011) also found that mothers‟ age at first birth has a significant effect on infant and child mortality showing that a child born to a younger mother experienced the highest risk of dying. These findings are consistent with the findings of (Abimbola et al (2012) and Pandey et al. (2012)).

According to the results, mother‟s current breastfeeding status was a significant determinant of under-five mortality showing that children born to non-breastfeeding mother experience higher risk of mortality than children born to breastfeeding mothers. This result is consistent with the findings by (Mustafa and Odimegwu (2008), Pandey et al. (2012)).

The findings suggested that under-five child mortality risk is higher for children of poor mothers compared to children of medium and rich mothers. This finding is consistent with Mustafa and Odimegwu (2008), Ettarh and Kimani (2011), Wang (2003) and Ermias (2013).

The study indicated that children born from working mothers have higher risk of mortality than non-working mothers. This finding is consistent with Ermias (2013).

The results of ZINB indicated that marital status has a significant factor on the odds of being in the always zero group. The odds of being in the always zero group for unmarried mothers were higher than unmarried mothers.

This study has also identified the risk factors of under-five mortality in urban Ethiopia based on EDHS 2011 data. The total number of urban women covered in the present study was 1,807 of which 18.2% of them experienced under-five deaths due to different factors.

Children from urban areas, whose mothers are in a relatively older age group, were better in their survival status as compared to children whose mothers are in the younger age group. This study revealed that children born to very young (age less than 15years) mothers are more likely to die.

This finding is consistent with the findings of (Amare et al (2007); Ettarh and Kimani (2011); Gideon (2012)).

Level of education of Mothers in urban areas showed a significant effect on infant and child mortality. Under-five mortality decreased with increased level of mother‟s education. This result is consistent with the findings of (Frank et al (2001); Gideon (2012); Amare et al (2007)).

The under-five mortality risk was lower for women with access to toilet facility than that for women without toilet facility. This finding is consistent with Gedion (2012) and Bariagaber(2013). In addition, children born from working mothers have higher risk of mortality than non-working mothers.

The results of ZIP indicated that region and marital status have a significant factors on the odds of being in the always zero groups.

The infant and child mortality widely varied between regions in Ethiopia. The expected number of under-five deaths for women from Oromiya, Benshangul-Gumuz, SNNP and Gambela were higher than the expected number of under-five deaths for women in Tigray. But, the expected number of under-five mortality in Addis Ababa was lower than that of Tigray. This finding is confirmed by the previous researches (Tibebu (2011); Desta (2011); Chowdhury (2013)).

The ZINB regression analysis confirmed that mother‟s education is a significant determinant of child mortality showing that children born to illiterate mother experience higher risk of mortality than children born to mothers with primary and above education. This finding is consistent with the findings by Desta (2011),Tibebu (2011) Senayit (2012), Tesfaye (2011), Abdullah (2014), Mohammed (2013), Bogaarts (2008), Kasahun and Teshome (2014), Van Raalte, Kunst, et al (2012), Lemani (2013), Goro (2007), Kyei (2011), Hill (2001), Worku (2011), Buor (2002), Rafiqul et al (2013), Matthias et al (2012), Wang (2003), Buwembo (2010), Iram and Butt (2008), Gebremariam (2001), mutunga (2007), Ezra and Gurum (2002) and Anderson et al (2002).

When we look at the age of mothers at first birth, the expected number of under-five deaths for mothers in the age groups 15-19 and 20 and above years is less than that for mothers under 15 years of age. This study found that children born to very young (aged less than 15 years) mothers are more likely to die and these findings are consistent with Desta (2011), Lemani (2013) , Chowdhury (2013), Mondal et al.(2009), Tarihu and Eshetu (2013), Reuben (2013), Tesfaye (2011), Mohammed (2013), Kasahun and Teshome (2014), Ezra and Gurum (2002) and Mutunga (2007) findings.

In addition, the infant and child mortality was found to be significantly lower for the children whose mothers were breastfeeding than children whose mothers were non-breastfeeding. This result is confirmed by the previous researches (Chowdhury (2013); Kyei (2011); Bello & Joseph (2014); Uddin et al. (2009); Worku (2011); Sampson (2014); Tesfaye (2011)). Moreover, the expected number of under-five deaths for working mothers was higher than that for non-working mothers. This is consistent with the results of other researchers (Tibebu (2011), Kyei (2011), Mohammed (2013), Chowdhury et al. (2010), Rafiqul et al (2013), Uddin et al. (2009) and Iram and Butt (2008).

According to the results, under-five mortality risk is higher for children of poor mothers compared to children of medium and rich mothers. This finding is consistent with Tibebu (2011) and Abdullah (2014).

In addition, from the two categories of place of residence, woman in rural areas were less likely to be in the zero group compared to women in urban areas. Moreover, the odds of being in the always zero group for unmarried women were higher than the odds for married women.

Chapter Five

Conclusions and Recommendations

5.1 Conclusions

The main objective of the study was to identify some of the factors that influence the number of infant and child mortality not only at national level but also separately at rural and urban levels. The study was based on secondary data obtained from the central statistical agency of Ethiopia. Among the four models considered for analyzing the data from women in rural areas and at national level, the ZINB regression model was found to be the most appropriate model. The comparative analysis of the count data models suggested that the ZINB model is most appropriate to deal with the problem of over-dispersion and, at the same time, to model the inflation of zero values. Similarly, the ZIP regression model was found to be the most appropriate model for the data from women in urban areas.

Factors influencing the number of under-five deaths have been identified. The study revealed that mother‟s age at the first birth, currently breastfeeding, wealth index, current mother working, region and mother‟s level of education had statistically significant effect on the number of under-five deaths in rural parts of Ethiopia. Similarly, mother‟s level of education, age of mothers at the first birth, toilet facility and work/employment status of mothers were found to be statistically significant with the number of under-five deaths per mother in urban parts of Ethiopia. Also region, age of mothers at the first birth, mother‟s level of education, breastfeeding status of mothers, wealth index and employment status of mothers were found to be statistically significant with the number of under-five deaths in Ethiopia.

5.2 Recommendations

Based on our findings we recommend that

- There is a need for comprehensive prevention strategies that will help to further reduce child mortality.
- The government/ministry of health should give greater attention to improve immunization services and concentrate on health education campaigns for mothers and for the community.
- Early marriages should be discouraged and awareness about the danger of giving birth at early ages should be created through education.
- Health interventions should particularly be targeted towards women who are suffering from illness and weakness to allow them to continue breastfeeding.
- Effort should be made for providing better access to education and health facilities for mothers so that the gab in under-five mortality is bridged.

References

- Abdullah A. (2014). "Zero-inflated regression models for count data: an application to under-5 deaths".
- Abimbola O,Adepoju AO, Akanni O and Falusi AO. Determinants of child mortality in rural Nigeria. World Rural Observations [2012;4(2):38-45]. ISSN: 1944-6543 (Print); ISSN: 1944-6551. http://www.sciencepub.net/rural.
- Agresti,A. (2007), An introduction to Categorical data analysis, 2nd edition, John Wiley and sons; Hoboken, New Jesey.
- Akaike, H. (1973). "Information Theory and an Extension of the Maximum Likelihood Principle." Pp. 267–281 in *Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki. Budapest: Akademiai Kaido.
- Amare, D., Belaineh, G., and Fasil, T. (2007). Determinants of under-five mortality in Gilgel Gibe Field Research Center, Southwest Ethiopia, Ethiopian journal of health devlopment, vol. 21(2):117-124.
- Anderson,B.A.,Romani,J.H.,Phillips,H.E.and Van Zyl,JA.,20002.‟ Environment, access to health care and other factors affecting infant and child survival among the African and coloured populations of South Africa, 1989-1994‟. Population and Environment Vol.23 No. 4.
- Antai, D. (2011) Regional inequalities in under-5 mortality in Nigeria: a population-based analysis of individual- and community –level determinants. Population Health Metrics, 9, 6. doi: 10.1186/1478-7954-9-6.
- Bariagaber, H. (2013) .Housing Correlates of Under-Five1 Mortality in Urban Ethiopia: A Paper Presented at the XXVII IUSSP International Conference, Busan, Korea, 26-31 August 2013. Department of Population Studies, P/Bag UB 00705.
- Bello, R. A. & Joseph, A. I. (2014). Determinants of Child Mortality in Oyo State, Nigeria. Africa research review, An International *Multidisciplinary Journal, Ethiopia Vol. 8 (1), Serial No. 32, January, 2014:252-272.*
- Bongaarts, J. Fertility Transitions in Developing Countries: Progress or Stagnation? Working Paper Number 7 of 2008. Population Council, New York. pp 2-19. [Accessed: 19 March 2012].

- Buor, Daniel. (2002). Mothers' education and childhood mortality in Ghana. Health Policy, 64, 297-309.
- Buwembo, p. (2010).Factors associated with under-five mortality in South Africa: trends 1997-2002.
- Chowdhury,H.A (2013).Determinants of Under-Five Mortality in Bangladesh. Journal of statistics, Vol. 3, pp 213-219.Available from http://dx.doi.org/10.4236/ojs.2013.33024.
- Chowdhury Q.,Islam andR.,Hossain K., (2010). "Socio-economic determinants of neonatal, post neonatal, infant and child mortality in Bangladesh". International Journal of Sociology and Anthropology Vol. 2(6), pp. 118-125.
- Dalrymple, M.L., Hudson, I.L., Ford, R.P.K., "Finite Mixture, Zero-Inflated Poisson and Hurdle Models with Application to SIDS", Computational Statistics & Data Analysis, 41: 491-504 (2003).
- 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, Lund University.
- Ermias Dessie Buli(2013). Determinants of Child Survival Chances in Rural Ethiopia. Adama Science and Technology University, Adama, Ethiopia. Proceedings 59th ISI World Statistics Congress, 25-30 August 2013, Hong Kong (Session CPS105).
- Espo, M. (2002). Infant Mortality and its underlying Determinants in Rural Malawi (Dissertation). University of Tampere Medical School.
- Ethiopia Demographic and Health Survey, Central Statistical Authority, Addis Ababa, 2011.
- Ettarh,R. and Kimani,J. (2011). Determinants of under-five mortality in rural and urban Kenya. African Population and Health Research Center, Nairobi, Kenya. Rural and Remote Health 12: 1812. Available: http://www.rrh.org.au.
- Ezra. M. and Gurum, E. (2002). Breastfeeding, birth intervals and child survival: analysis of the 1997 community and family survey data in southern Ethiopia.
- Filmer D and Pritchett, L. 2010. Child Mortality and Public Spending on Health: How Much Does Money Matter? Policy Research Working Paper No. 1864. The World Bank Washington D.C.

- Fitsum Z. (2010) „Socio-economic factors of early child hood mortality in Ethiopia, working paper No 3/2010.
- Frank H., Mekonen T., Paolo V. (2001): Infant and child mortality in urban Ethiopia: with special reference to socio-demographic and housing conditions in urban areas and Addis Ababa. Italian Multi-Bi Research Project ETH/92/P01. Institute for Population Research – National Research Council (Irp-Cnr) Roma, Italy.
- Gakusi, E. and M. Garenne. 2006. "Health transitions in Sub-Saharan Africa: overview of mortality trends in children under five years old (1950- 2000)", Bulletin of the World Health Organization 84(6):470-478.
- Gardner, Howard (1995). Reflections on multiple intelligences. Phi Delta Kappan, 77 200-Generalized Linear Models (with J.A. Nelder), (1983), Chapman and Hall, London.
- Gebremariam, W. (2001) „Diarrheal morbidity among young children in Eritrea: environmental and socioeconomic determinants", Journal of Health Population Nutrition 19(2): 83-90.
- Gideon R., (2000). "Analysis of regional differentials in under-five mortality in Kenya using a count-data regression model", ACAP Working Paper No 14.
- Gideon R., (2012). Under-five mortality differentials in urban East Africa: a study of three capital cities (Nairobi, Dar-es-Salaam and Kampala). *African Population Studies Vol* 26, 1.
- Goro, M. (2007). The stalling child mortality: the case of three northern regions. The 5th Conference of Union for African population, Tanzania.
- Greene, W.H. 1994. Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. Working paper No. 94-10, Department of Econometrics, Stern School of Business, New York University, New York, U.S.A.
- Hammer GP, Some F, Muller O, Kynast-Wolf G, Kouyate B, & Becher H. (2006). Pattern of cause-specific childhood mortality in a malaria endemic area of Burkina Faso. Malaria Journal, Online edition.
- Hilbe, J.M., 2007. Negative Binomial Regression, Cambridge University Press.
- Hill, K., Bicego, J. & Mahy, M. (2001). Childhood mortality in Kenya: An examination of trends and determinants in the late 1980s to mid 1990s..

- Iram, U. and Butt, M. S. (2008) „Socioeconomic Determinants of Child Mortality in Pakistan: Evidence from Sequential Probit Model", International Journal of Social Economics, Vol. 35 No. 1/2, 2008.
- Jacoby, H. & Wang, L. (2003). Environmental Determinants of Child Mortality in Rural China: A Competing Risks Approach. Washington DC: World Bank.
- Jang, T. Y. (2005). Count data models for trip generation. Journal of Transportation Engineering, 131, 444-450.
- Kanmiki, E.W**,** Bawah, A.A**,** Agorinya, I.**,** Achana**, F.S,** Awoonor-williams**,J. K,** Oduro, A.R**,** Phillips, J.F. and Akazili, J. (2014). Socio-economic and demographic determinants of under-five mortality in rural northern Ghana. *BMC International Health and Human Rights* 2014, 14:24. doi:10.1186/1472-698X-14-24.
- Kasahun T. and Teshome K., (2014). " Correlation of under-five child mortality in Ethiopia".
- Kilobi, J. (2009): Factors influencing under -five mortality in urban and rural Kenya.
- Klaauw, van der B. & Wang, L. (2003). Child Mortality in Rural India. World Bank Working Paper, Washington DC: World Bank.
- Kouame P., (2014), "determinants of regional disparities in under age five mortality in Cote d'Ivoire".
- Kyei, K. A. (2011). Socio - Economic Factors Affecting under five mortality in South Africa- An Investigative Study. *Emerging Trends in Economics and Management Sciences*, 2(2): 104-110.
- Lambert, D., "Zero-inflated Poisson regression, with an application to defects in Manufacturing", Technometrics, 1992; 34:1–14.
- Lemani, C. (2013): Modeling Covariates of Infant and Child Mortality in Malawi, University of Cape Town.
- Leslie, Ofosu Tettey 2003. "Socio-economic Differentials of Infant and Child Mortality in Ghana and Nigeria". M.A. Thesis RIPS, University of Ghana, Legon.
- Matthias M. Okposio, Monday O. Unior and Felix O. Ukpeteru, (2012). "Socio-demographic determinants of mortality in hospitalized under-five children at a secondary health care centre in the Niger Delta". International Journal of TROPICAL DISEASE & Health 2(3): 173-181, 2012.

- Mesike, C.G.and Mojekwu J. N. (2012), Environmental Determinants of Child Mortality in Nigeria. Journal of Sustainable Development, 5(1): 65-75.
- Min,Y., Agresti, A. (2005). Random effects models for repeated measures of zero-inflated count data. Statistical Modelling, pp. 1-19.
- Mohammed A., (2013). "Supportive strategies regarding accidents prevention for mothers of children under five years old", Journal of Biology, Agriculture and Healthcare, Vol.3, No.20, 2013.
- Mondal, N.I., Hossain, K. and Ali, K. (2009). Factors Influencing Infant and Child Mortality: A Case Study of Rajshahi District, Bangladesh, *J Hum Ecol*, 26(1): 31-39.
- Mosley, W. and Chen, L. (1984). An analytical framework for the study of child survival in developing countries. Population and Development Review vol.10: 25-45.
- Mustafa, E. & Odimegwu, C. (2008). Socioeconomic determinants of infant mortality in Kenya: Analysis of Kenya DHS 2003.
- Mulugeta, F. Z. (2012) „Socioeconomic Factors Affecting Childhood Mortality In Ethiopia: An Instrumental Variable Approach", *Ethiopian Journal of Economics* XX(2):63-81.
- Mutunga, C. J. (2004). Environmental determinants of child mortality in Kenya. Kenya Institute for Public Policy Research and Analysis (KIPPRA), Nairobi, Kenya.
- Mutunga, C. J. (2007). Environmental Determinants of Child Mortality in Kenya, UNU-WIDER Research paper No. 2007/83. Helsinki: United Nations University World Institute for Development Economics Research.
- Nelder, J .A. and Wedderburn R.W.M., (1972), Generalized linear models, Journal of the statistical society Cameron, A.C., Trivedi, P.K., (1998), Regression Analysis of Count Data, Cambridge University Press, Cambridge, U.K.
- Pandey, R.M, Mani, K. and Dwivedi, S.N. (2012). Determinants of Under-Five Mortality in Rural Empowered Action Group States in India: An Application of Cox Frailty Model. International Journal of MCH and AIDS (2012), Volume 1, Issue 1, pp 60-72.
- Pedan, Alex. "Analysis of Count Data Using the SAS System." *SUGI* 26, (2001): 22-25.
- Rafiqul Islam, Mahfuzar Rahman, Obaidur Rahman, (2013) a study on infant and under-five mortality in Bangladesh: life table approach. Science and Technology 2013, 3(3): 74-84.

- Raftery, A. E. 1986. “Choosing models for cross-classifications.” American Sociological Review 51:145–146.
- Reuben, M. (2012): Determinants of under-five mortality in Kenya during upsurge and declining period.
- Sampson, B. 2014.Determinants of under-five mortality in the Tano South district of Ghana, M.SC. Thesis, Kwame Nkrumah University, Kwame Nkrumah.
- Senayit, S (2012): Determinants of mortality among one to five Years old children in Ethiopia, Addis Ababa University.
- Sturman, M. C. (1999). Multiple approaches to analyzing count data in studies of individual differences: The propensity for type I errors, illustrated with the case of absenteeism prediction. Educational and Psychological Measurement, 59:414–430.
- Tarihu,D. and Eshetu,G. Social determinants of under-five mortality in Ethiopia: Event history analysis using evidence from Ethiopian Demographic and Health Survey (EDHS).Vol.5, No.5, 879-884 (2013). http://dx.doi.org/10.4236/health.2013.55115.
- Tesfaye A., (2011). “Modeling children of under-five mortality in Ethiopia using Bayesian approach”.
- Tibebu, G. (2011): Identification of risk factors and regional differentials in under-five mortality in Ethiopia using multilevel count model, Addis Ababa University.
- Uddin, J., Hossian, Z., Ohid, M. (2009). Child mortality in a developing country: a statistical analysis, Journal of applied quantitative methods, Vol.4 (3):270-276.
- UN (1995): Women‟s education and fertility behavior: recent evidence from the demographic health survey, New York, United Nations.
- UN (2014): Levels and trends of in child mortality reports 2014.
- UNICF (2009): States of the world‟s children. (Accessed on 16 oct 2010).
- UNICF (2012): Trends and levels in child mortality reports 2012: The United Nations inter-agency groups for child mortality estimates.
- United Nations Children‟s Fund, the state of world‟s children 2007: women and children-double dividend of gender equality, UNICEF, New York, Dec 2008.
- 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.

- Vuong QH (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica, 307–334
- Wang, L. (2003). Environmental Determinants of Child Mortality: Empirical Results from the 2000 Ethiopia DHS. Washington DC: World Bank.
- Wang, P., Puterman, M.L., Cockburn, I.M., Le, N., “Mixed Poisson Regression Models with Covariate Dependent Rates”, Biometrics, 52: 381-400 (1996).
- Wang, Z. 2011. One mixed negative binomial distribution with application. Statistical Planning and Inference. 141, 1153-1160.
- WHO (2014): Levels and trends of in child mortality, New York, Sep 2014.
- Worku, Z.B. (2011). A survival analysis of South African children under the age of five years. *Health SA Gesondheid,* 16(1): 90-101.
- World Health Organization, United Nations Children‟s Fund, United Nations Population Fund and the World Bank, Trends in Maternal Mortality: 1990-2010,Estimates developed by WHO, UNICEF, UNFPA and the World Bank, 2012.
- Yared, S. (2013): Survival analysis of under-five mortality in rural parts of Ethiopia, Addis Ababa University.

Appendix

Table A1: Frequency of number of under-5 deaths in rural parts of Ethiopia

Number of deaths per mother	Frequency	percent
0	5591	64.50
1	1802	20.8
2	772	8.9
3	303	3.5
4	131	1.5
5	42	0.5
6	15	0.2
7	5	0.1
8	4	0.0
9	1	0.0
10	1	0.0
15	1	0.0
Total	8668	100.0