Addis Ababa University

College of Natural and computational Sciences

Department of Statistics

Determinants of under-five child mortality in Ethiopia: A spatial and multilevel count regression analysis.

BY

Gebremedhin Melkie

Advisor Mekonnen Tadesse (Assoc Prof.)

A thesis submitted to the Department of Statistics in partial fulfillment of the requirements for the Degree of Master of Science in Statistics (Biostatistics)

June, 2024

Addis Ababa, Ethiopia

List of Acronyms

GTP	Growth and transformation plan
SDG	Sustainable development goals
WHO	World Health Organization
EMDHS	Ethiopian mini demographic and health survey
ZIP	Zero inflated Poisson
ZINB	Zero inflated negative binomial
U5CM	Under-five child mortality
PHC	Population and housing census
EAS	Enumeration areas
EPHI	Ethiopian public health institute
ICF	Inner city fund
GIS	Geographic information system
GLMs	Generalized least square method
NB	Negative binomial
MLE	Maximum likelihood estimate
LRT	Likelihood ratio test
AIC	akakie information criteria
BIC	Bayesian information criteria
MZINB	Multilevel zero-inflated negative binomial
ICC	Intra class correlation
IRR	Incidence rate ratio

Declaration

This is to certify that the thesis entitled "Determinants of under-five child mortality in Ethiopia: A spatial and multilevel count regression analysis using 2019 Ethiopian mini demographic and health survey data" submitted in partial fulfillment of the requirements for the degree of Master of Science in biostatistics at Addis Ababa university, is a record of original work carried out by me and has never been submitted to this or any other institution to get any other degree or certificates. The assistance and help I received during the course of this investigation have been duly acknowledged.

____________________ ____________________

Name of the candidate signature

Department of Statistics

College of Natural and Computational Sciences

Addis Ababa University

This is to certify that the thesis prepared by Gebremedhin Melkie entitled: " Determinants of under-five child mortality in Ethiopia: A spatial and multilevel count regression analysis.” submitted in partial fulfillment of the requirements for the Degree of Master of Science in Statistics (Biostatistics) complies 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

Acknowledgments

I wish to express my utmost gratitude to the Divine Creator, GOD Almighty, for His boundless grace and mercy that have sustained me thus far. Next, I thank the Virgin Mary for giving me the good journey during the long night of my life. I express my sincere gratitude to my advisor Mekonnen Tadesse (Assoc Prof.) for his time, continuous support, patience, motivation and immense knowledge towards the writing of this thesis. His matured experience and knowledge guided me in all the time of research and writing of this thesis. And also my sincere thanks to all other lecturers in the Statistics department who provided me with valuable inputs regarding my work and helped me all along. Furthermore, I would like to thank colleague and friends that contributed in diverse ways to my success in this program; I say a big thank you. Finally, my thanks go to my family, for their continued support, prayer, contributions and bearing with me throughout this program of study.

Abstract

***Introduction**:- The under-five mortality rate serves as a significant indicator of child health and overall development within countries, highlighting the quality of life within a given population.*

Tragically, millions of children below the age of five lose their lives each year. In Ethiopia, the burden of under-five deaths remains unevenly distributed.

***Objective:-** This study aims to examine the spatial distribution of under-five mortality in Ethiopia and identify factors impacting number of under-five child deaths using spatial analysis methods and multilevel count regression models.*

***Materials and Method:-** the data was retrieved from the EMDHS 2019, and the total unweighted number of 5753 women were used in this study. Various count models were considered and the multilevel zero inflated negative binomial model happened to be the best model to fit under-five child mortality data. IRR (incidence rate ratio) has been used for interpretations. Hotspot and cold spot, as well as kriging interpolation analysis were done for describing the spatial variation of under-five child mortality.*

***Result:-** Spatial patterns in the number of under-five deaths were observed. The global Moran's I value (I = 0.416267, z score =5.410983, P value < 0.001) revealed a substantial group of under-five child mortality throughout the country. Under-five death had a statistically significant regional clustering pattern among Ethiopian children. The multilevel zero inflated negative binomial model fit results revealed that the expected number of under-five child mortality in female headed households is 1.27 times higher than the expected number of under-five deaths in male headed households (IRR=1.27). Also, the expected number of under-five deaths in rural areas is 0.6356 times the expected number of under-five deaths in urban areas. Likewise, when water is accessible on the premises, the number of under-five deaths is 0.05779 times lower compared to where obtaining water takes more than 30 minutes. And the expected number of under-five deaths when the mother uses contraceptive methods decreased by a factor of 0.79 (IRR=0.79) compared to the expected number of under-five deaths when the mother is not using contraception. Moreover, the expected number of under-five deaths for a birth order of 2^{nd} and 3^{rd}, 4^{th} and 5^{th}, and 6th and above is 2.10, 4.0812 and 9.73 times the expected number of under-five deaths for a first birth respectively. Moreover, multiple birth increases the incidence rate of under-five mortality (IRR=1.5595).*

The spatial analysis verified hot and cold spot areas of under-five mortality among under-five children in Ethiopia

***Conclusion:-** The spatial analysis verified hot(high risky) and cold (less risky) spot of under-five mortality among under-five children in Ethiopia. The zero inflated negative binomial model revealed that under-five deaths per mother differ in age at first birth time, current contraceptive use, birth order, and type of birth. Place of residence, time to get water, household size, sex of household head and age of household head are statistically significant factors. Policy makers should develop targeted policies that consider the specific challenges and needs of different regions to address geographical variations in under-five mortality.*

***Keywords:-** under-five mortality, count models, spatial analysis, multilevel analysis.*

List of tables

List of figures

Table of Contents

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND

Mortality refers to the state or condition of being mortal. Mortality is also the number of deaths within a particular society and within a particular period of time (Wikipedia, 2024).

Historically, mortality has often been used as a barometer of welfare. One of the demographic variables that affect population trends is mortality (Ayele, T Zewotir, 2016).

The under-five mortality rate is represents the likelihood of a child born in a particular year or timeframe dying before reaching the age of five, subject to the age specific mortality rates of the period (WHO, 2019).

Preventing the life of the newborn and reducing the entrenched disparity of childhood mortality across different levels is one of the crucial public health problems, especially in underdeveloped and developing countries in the world. Sustainable development goal (SDGs)-3.2 is aimed at terminating all preventable under-five child mortality and shrinking it to 25 per 1000 live births or lower than this by 2030. Several factors have been shown to be linked with childhood mortality Bhusal and Khanal (2022).

Child mortality rates reflect a country‘s level of the social-economic development and quality of life and are used for monitoring and evaluating population, health programs and policies. Child mortality rates are unacceptably high in many developing countries and need to remain the focus of public policy to gain improvement in infant and child survival. The under-five mortality rate is a leading indicator of the level of child health and the overall development of countries. It indicates the quality of life of a given population, as measured by life expectancy and understanding the determinant factors of under-five mortality is essential to inform public health policies and design strategies to accelerate the reduction of under-five mortality. (GA Kaplan, JE Keil , 1993).

The age at death of children under the age of five refers to the period or age at which they pass away before celebrating their fifth birthday. This demographic events, known as death, has a significant impact on population development. The mortality of children is closely monitored by demographers, policymakers, and researchers because it serves as crucial indicator of a countries' health condition. Globally, under-five death accounts approximately 5.2 million fatality, with around 47% occurring within the first month of life, 28% between 1 and 11 months and 25% between 1 and 45 years of age (Agidew, Belay, 2023)

According to the world Health Organization (WHO), in 2020 an estimated 5 million children under the age of 5 years were died, mostly from preventable and treatable causes. About 80% of these occurs in sub-Saharan Africa and south Asia, and just six countries account for half of all under-five deaths: china, India, Pakistan, Nigeria, Ethiopia and Democratic Republic of the Congo.

In almost half of the cases, malnutrition plays a role, while unsafe water, inadequate sanitation and hygiene are also significant contributing factors. The world has made substantial progress in reducing child mortality in the past few decades. Estimates Developed by the UN Inter-agency Group for Child Mortality Estimation, (Keeley and Little 2017) shows that the global under- five mortality rate declined by 56 % from 93 deaths per 1,000 live births in1990 to 41 in 2016. Progress in reducing child mortality has been accelerated during the 2000– 2016 period compared with the 1990s. That is, the annual rate of reduction in the under- five mortality rate has increased from 1.9% during 1990–2000 to 4.0 % during 2000– 2016. In 2016, an estimated 5.6 million children died before their fifth birthday of which 2.6 million (46 %) died in the first month of life ((Bizzego, 2021)

Globally, under-five mortality has dropped significantly by almost 45 percent between 2009 and 2011, but this progress is not the reality for all countries, despite much progress in advanced countries. In recent years, however, under-five mortality rate has declined and even reversed in many sub-Saharan African countries while they have continued to improve in other regions. The under-five mortality rate (U5MR) is generally 29 times higher in developing nations. The global under-five mortality rate declined by 59 percent; from 93 deaths per 1,000 live births in 1990 to 38 in 2021.

Compared to developed countries, approximately 6.3 million infants and children under five years of age die each year, with large variations in under-five mortality rates and trends across regions and countries. A 2013 report on child mortality states that about three-quarters of all child deaths happened in two regions: Africa (46%) and Southeast Asia (28%). More than 50% of these deaths were clustered in only six countries: China, Democratic Republic of the Congo, Ethiopia, India, Nigeria, and Pakistan. On average, 1 out of every 11 children born in sub-Saharan Africa dies before age five. This is nearly 15 times the average rate (1 in 159) in high-income countries. The mortality rate of children under five in Ethiopia has been declining. The decline of the role of agriculture in the national economy, the increase of urbanization, and the launching of globalization have accelerated the economic performance of the country and significantly changed the trend of mortality rates particularly the mortality rate of children under five. The trend in demographic and health indicators from 2000 to 2016 EDHS shows improvement of child health status. Early childhood mortality (all neonatal mortality, postnatal mortality, infant mortality, child mortality, and under-five mortality) has been decreasing in Ethiopia. For example, under-five mortality has decreased from 123 deaths per 1000 live births in 2005 to 88 deaths per 1000 live births in 2011, 67 deaths per 1000 live births in 2016 and 59 deaths per 1,000 live births in 2019. Still, Ethiopia is at a lower position than other East African countries in terms of many child health indicators (EPHI and ICF, 2021).

1.2 Statements of the Problem

Child mortality is a significant focus within the Sustainable Development Goals (SDGs) aimed at achieving global development. However, under-five deaths continue to be disproportionately distributed. Approximately, 80% of these deaths occur in Sub-Saharan Africa and Asia (Fenta, Ayenew, 2020).

Every year, more than 257,000 children under the age of five die and 120,000 die in the neonatal period (USAID, 2019). If the situations continue as such, more than 3,084,000 children will die until 2030. This requires a comprehensive and thorough study of the determinant of under-five mortality in the country in order to find out and suggest strategies to help improve the current situation. In Ethiopia, there have been regional variations in the number of under-five mortality (Fenta, Ayenew, 2020).

In Ethiopia, there have been regional variations in number of under-five mortality (Fenta, Ayenew, 2020). In this study, region is assumed to have an effect on modeling the determinants of number of under-five mortality, which may be due to the heterogeneity in regions of the study. The under-five mortality was and is now a major public health problem in Ethiopia. The prevalence of under-five mortality and socioeconomic and socio-demographic factors that affect under-five mortality have been assessed. However, the gap those studies is the variables considered are not well addressed and their spatial patterns have not been well described.

Moreover, most of the previous studies used binary logistic regression and survival analysis without showing that under-five mortality varies across different physical, ecological, and political structures within countries and without considering the number of deaths. Many of the studies dichotomized the count variable and performed the analysis, losing information. The logistic regression cannot be used to identify relevant risk factors of the number of under-five deaths (Guddattu, V., Rao, K., & Rajkannan, T, 2015).

The study is expected to answer the following questions.

- What are the factors that have significant impact on the number of under-five child mortality in Ethiopia?
- What is the spatial pattern of under-five child mortality in Ethiopia?
- Which count regression model is better to analyze under-five child mortality?

1.3 Objectives of the Study

1.3.1 General Objectives

The general objectives of this study was to investigate the spatial distribution of under-five mortality and identify the factors that affect the number of under-five child mortality in Ethiopia using spatial and multilevel count regression models.

1.3.2 Specific Objective

- To determine significant factors associated with the number of under-five child mortality in Ethiopia.
- To analyze the spatial pattern of under-five child mortality and identify the area that has high risk of child mortality in Ethiopia.
- Measure the variability of under-five child mortality by individual and community level predictors.

1.4 Significance of the study

The findings of the study is expected to help in formulating appropriate health programs and policies that will help to meet the United Nations Sustainable Development Goal (SDG) of reducing under-five mortality to at least as low as 25 deaths per 1,000 live births by 2030 (Unicef, 2017). It can also be used to take more cost-effective interventions and policies to reducing child mortality and to improving the health and life expectancy of the society.

- ✓ Help stakeholders in the planning, formulation and implementation of policies concerning the reduction of under-five mortality.
- ✓ Provide additional information to policy makers and researchers that can be used for further studies on under-five child mortality.

1.5 Limitation of study

The study used data from the EMDHS 2019, so the results may not accurately represent the current situation of under-five child mortality in Ethiopia.

CHAPTER TWO

LITERATURE REVIEW

2.1 Over View of Under-Five Mortality

Under-five mortality has continued as a key challenge to public health in Ethiopia, and other sub-Saharan African countries. The threat of under-five mortality is never-ending and more studies are needed to generate new scientific evidence (Oloo, 2005).

Ethiopia has made remarkable strides in reducing child mortality over the past two decades. However, the reduction of under-five mortality remains a significant concern in light of the Sustainable Development Goals (SDGs) introducing in 2016. To achieve the Sustainable Development Goals (SDGs), Africa as a whole would need to achieve an average annual decline rate of approximately 8% between 2015 and 2030. Despite a global decline in under-five mortality rates, the risk of children dying before their fifth birthday remains highest in African countries, particularly in sub-Saharan Africa.

Recent evidence indicates that countries in the Saharan region exhibit alarmingly high Under-five Mortality Rates (U5MR), reaching 78 deaths per 1000 live births. To put this into perspective, it means that approximately 1 in every 13 children dies after surviving the first month but before their fifth birthday. This ratio is 16 times higher than the average rate of 1 in 199 observed in high-income countries. Notably, Ethiopia is one of the five countries that collectively account for half of all global under-five deaths in 2018, with a U5MR of 67 deaths per 1000 live births. Despite the implementation of progressive and consistent health interventions, Ethiopia continues to experience an upward trend in U5MR, signifying the need for sustained efforts to address this critical issue (Liyew, 2021).

This indicator is influenced both directly and indirectly by a number of factors. Because the immune system of young children is not as developed as that of a healthy adult, children are more likely to be affected by the availability of health facilities, nutritious food, and sanitation, or lack thereof. In developing countries, child mortality accounts for a relatively higher proportion of all deaths, whereas in the developed countries, it represents an increasingly small segment of total mortality (Oloo, 2005).

According to the World Health Organization (WHO), significant progress has been made in reducing under-five child mortality rates worldwide. However, disparities persist, with Sub-Saharan Africa and South Asia accounting for the majority of child deaths. A study by Liu et al. (2020) analyzed global trends from 1990 to 2019 and emphasized the importance of addressing preventable causes such as neonatal conditions, pneumonia, diarrhea, and malaria (Liu L, 2020).

Understanding the causes and risk factors associated with under-five child mortality is crucial for effective interventions. A study by (Amouzou A, 2018)examined data from 51 countries and identified the leading causes of child mortality, including preterm birth complications, pneumonia, and diarrhea. The study also highlighted the importance of addressing risk factors such as low birth weight and lack of access to quality healthcare.

2.2 Empirical Study

Berhanu Teshome utilized the 2016 Ethiopian Demographic and Health Survey dataset to determine the effect of underlying socioeconomic, demographic, and cultural factors on under-five mortality in Ethiopia by using multivariate logistic regression model. It was found in this study that the risk of under-five mortality has significant associations with the region area of residence, educational level of the mother, current age of mother, mother's age at first birth, occupation of mother, partner's occupation, birth order, preceding birth interval, and birth type. The study explicitly shows the existence of inconsistency in the distribution of under-five mortalities among the regions of Ethiopia. Ethiopia has achieved significant declines in under-five and infant mortality rates (Woldeamanuel, 2019).

Reducing under-five child mortality in Ethiopia represents a significant and formidable challenge that requires urgent attention. The Millennium Development Goal (MDG) 4 set out to achieve a two-thirds reduction in under-five mortality, yet this objective remains unfulfilled for the majority of middle and lower-income countries. The task of reaching the target of 55 deaths per 1000 live births for under-five mortality by 2030 is undoubtedly difficult. Moreover, studies have uncovered substantial regional disparities, highlighting notable variations in under-five mortality rates across different areas (Worku, 2021).

Spatial clustering of under-five mortality was observed in Ethiopia. Over the past two decades, the Benishangul-Gumuz region consistently exhibited a higher risk of under-five mortality. Furthermore, additional hotspot areas were identified in Afar and Amhara (in 2000, 2005, and 2016), Gambela (in 2011), and the South Nation Nationality and People's (SNNP) Region (in 2016). These findings highlight specific regions within Ethiopia where under-five mortality rates are elevated and warrant targeted interventions to address the issue (Liyew, 2021).

Still under-five mortality remains a public health problem in Ethiopia. Being a female household head, age of mother at first giving birth, being employed, having multiple births, and childhood diarrhea were associated with a higher incidence of under-five mortality. (Geremew , 2020).

A study done on factors affecting child mortality in Pakistan by using binary logistic regression analysis, revealed that region, education of mother, birth order, preceding birth interval (the period between the previous child birth and the index child birth), breastfeeding , family size are significant factors associated with child mortality in Pakistan. Child mortality decreased as level of mother's education, preceding birth interval and family size increased. Child mortality was low for low birth orders. Child survival was significantly higher for children who were breastfed as compared to those who were not (Ahmed, Kamal, 2016).

A significant heterogeneity was observed between clusters for institutional delivery which explains about 57% of the total variation. A binary logistic regression was used to analyze child mortality (Awol M, 2023)

In Ethiopia, the death of children under the age of five has decreased in the last two decades, but it is still unacceptable with significant differences at sub-national and local levels. In the West, North East, East and North Central regions, the highest number of under-five deaths was observed. Population density is negatively associated with under-five mortality in Ethiopia. This study investigated the spatiotemporal distributions of under-five mortality in Ethiopia between 2000 and 2019 considering sub-national and local levels as well as several ecological-level variables (like temperature, precipitation...) related to child mortality (Atalell KA, Alene KA, 2023).

A study examining the factors influencing child mortality in Pakistan utilized binary logistic regression analysis. The findings revealed that several variables were significantly associated with child mortality. These included the region, education level of the mother, birth order, preceding birth interval (the time between the previous and index child's births), breastfeeding, and family size (Ahmed, Kamal, 2016).

The study found that child mortality rates decreased as the level of the mother's education, preceding birth interval, and family size increased. Additionally, children with lower birth orders tended to have lower child mortality rates. Notably, child survival rates were significantly higher among children who were breastfed compared to those who were not. These findings emphasize the importance of considering these factors in efforts to reduce child mortality in Pakistan (Ahmed, Kamal, 2016).

The generalized geo-additive mixed-effects model results identified several key factors associated with the composite index of anthropometric failure (CIAF) among children including gender of the child, presence of comorbidity, size of child at birth, dietary diversity, birth type, place of residence, age of the child, parental level of education, wealth index, sanitation facilities, and media exposure (Fenta, 2021).

CHAPTER THREE

DATA AND METHODOLOGY

3.1 Study Design and Study Area

Ethiopia is a landlocked country in the horn of Africa. The country lies completely within the tropical latitude and is relatively compact, with similar north-south and east-west dimension. Ethiopia is found in the Horn of Africa at (3°-14° N and 33°-48°E). The country covers 1.1 million Square kilometers and it has a high central plateau that varies from 4550 meters above sea level down to the Afar depression to 110 meters below sea level.

Administratively, Ethiopia is divided into nine geographical regions and two administrative cities by the time of the survey. The sample for the 2019 EMDHS was designed to provide estimates of key indicators for the country as a whole, for urban and rural areas separately, and for each of the nine regions and the two administrative cities (EPHI and ICF, 2021)

3.2. Source of Data

Ethiopian mini demographic and health survey was selected in two stages and stratified each regions of Ethiopia into urban and rural areas. EMDHS was implemented by Ethiopian public health institute (EPHI). The data collection was conducted from March 21, 2019, to June 28, 2019 in Ethiopia.

The 2019 Ethiopia Mini Demographic and Health Survey (EMDHS) is the second Mini Demographic and Health Survey conducted in Ethiopia.

The 2019 EMDHS sample was stratified and selected in two stages. In the first stage, a total of 305 EAs (93 in urban areas and 212 in rural areas) were selected with probability proportional to EA size (based on the 2019 PHC frame) and with independent selection in each sampling stratum. A household listing operation was carried out in all selected EAs from January through April, 2019. The resulting lists of households served as a sampling frame for the selection of households in the second stage.

In the second stage of selection, a fixed number of 30 households per cluster was selected with an equal probability systematic selection from the newly created household listing. A total of 8663 households were interviewed with 2645 households from urban and 6018 from rural areas. This study will use a sample of 5753 unweighted data of women (EPHI and ICF, 2021).

3.3. Source Population

All women of reproductive age living in Ethiopia.

3.4. Study Population

The study participants are women and their children under five years of age.

3.4.1 Inclusion Criteria

Women of reproductive age have been included

3.4.2 Exclusion Criteria

Women under 15 years and over 49 years of age were excluded from the analysis.

3.5. Variables in the Study

3.5.1. Dependent Variables

The dependent variable of this study, Y_{ij}, the number of under-five children deaths that each mother had in her lifetime, where i=0,1… denotes the number of individual under-five children in a household and j = the enumeration areas (EAs)/clusters

3.5.2 Independent Variables

The explanatory variables of the number of under-five child mortality include individual and community level variables.

Individual level predictors are: highest educational level of mother, birth interval, sex of household head, age of household head, mother's age at first birth, number of household member, marital status of mother, time to get water, sex of the children, multiple birth, Current Contraceptive utilization, Birth season and order of birth.

Community level predictors are region, place of delivery, residence, Source of drinking water and Type of toilet facility.

3.6 Methods of Data Analysis

In this study spatial analysis and multilevel count regressions are used to describe the spatial distribution of under-five mortality and to identify the individual and community level determinant factors of U5CM in Ethiopia. The response variable of the study is the number of under-five children who died before the EMDHS 2019 survey. First, the study examine the spatial pattern of U5CM in Ethiopia, and finally, the study was attempt to identify and assess the associated factors of mortality among under-five children using an appropriate multilevel count regression model.

Different statistical techniques have been used in the analysis of collecting data; however, methods of data analysis depend on the nature of the variables incorporated in the study and the data type of the basic variables. Taking the above points into account, the study will use descriptive and inferential statistics.

3.7 Spatial Analysis

Spatial analysis involves analyzing data by considering their location or spatial relationships. It includes techniques for visualizing phenomena, assessing spatial autocorrelation and modelling spatial relationships (Agidew, Belay, 2023)

Spatial analysis utilizes spatial data to derive results based on the geography of target area (MM Fischer, 2011).

Spatial analysis refers to studying entities by examining, assessing, evaluating, and modeling spatial data features such as locations, attributes, and relationships that reveal data's geometric or geographic properties. It uses a variety of computational models, analytical techniques, and algorithmic approaches to assimilate geographic information and define its suitability for a target system (MM Fischer, 2011).

Fundamentally, spatial analysis derives insights from spatial data, which provides a numerical representation of any physical entity in a geographic coordinate system. It serves as a spatial unit of a map (MM Fischer, 2011).

Spatial analysis stands over the principle that there is some spatial component; absolute relative, or both in data. Indeed, at the beginning of the twentieth century, 80% of all data have already some kinds of spatial explanation (JA Tenedório, J Rocha, 2018).

Spatial analysis is any of the formal techniques which studies entities using their topological, geometric or geographic properties (https://en.wikipedia.org/wiki/Spatial_analysis, 2023).

3.7.1 The Concept of Spatial Dependence

Spatial dependence refers to the degree of spatial autocorrelation between independently measured values observed in geographical space.

The first law of geography states that "all things are related, but nearby things are more related than distant things''. Observations made at different locations may not be independent. For example, measurements made at nearby locations may be closer in value than measurements made at locations farther apart. This phenomenon is called spatial autocorrelation. Spatial autocorrelation measures the correlation of a variable with itself through space. Spatial autocorrelation can be positive or negative. Positive spatial autocorrelation occurs when similar values occur near one another. Negative spatial auto correlation occurs when dissimilar values occur near one another (JA Tenedório, J Rocha, 2018).

Spatial dependence is the spatial relationship of variables values or locations. It is measured as the existence of statistical dependence in allocation of random variables, each of which is associated with a difference geographical location. Spatial dependence is of importance in applications where it is reasonable to postulate the existence of corresponding set of random variable at locations that have not been included in the sample (https://en.wikipedia.org/wiki/Spatial_analysis, 2023).

There are two types of spatial autocorrelation measures: global measures and local measures. “Globally”, implies that the measure obtained refers to the dataset as a whole, whether it is a whole country, continent, or region. “Locally”, implies taking into consideration every polygon and getting a measure for each one of them (Getis, 2009)

3.7.1.1. Global Measures of Spatial Autocorrelation

Global spatial autocorrelation is a measure of the overall clustering of the data which provides one correlation statistic to summarize the whole study area. But if there is no global autocorrelation or no clustering in the whole area, one can look for clusters at a local level using the measure known as local spatial autocorrelation.

3.7.1.2 Local Indicators of Spatial Autocorrelation

Local measures are focused, that is, they usually assess the spatial autocorrelation associated with one particular spatial unit.

To observe if there is a local spatial cluster of high or low values, and identify the regions that contribute the most to the clustering (spatial autocorrelation), measures of local spatial autocorrelation such as local Getis and Ord (1992) statistics and the local indicator of spatial association (LISA) are used (Anselin, 1995).Often our interest lies not only in determining whether the data as a whole exhibit spatial autocorrelation, but also, in identifying the specific observations that exhibit spatial autocorrelation with their neighbors.

3.7.2 Spatial Autocorrelation Analysis

Spatial autocorrelation can be defined as a particular relationship between the spatial proximity among observational units and the numeric similarity among their values; positive spatial autocorrelation refers to situations in which the nearer the observational units, the more similar their values (and vice versa for its negative counterpart). Moreover, when neighboring locations share a certain amount of information then we can say that the spatial autocorrelation or spatial dependence exists (Getis, 2009).

3.8. Multilevel Regression Models

The term "multilevel" refers to the distinct levels or units of analysis, which usually, but not always, consists of, individuals (at a lower level) who are nested within contextual/aggregate units (at a higher level).

Multilevel models are statistical models which allow not only independent variables at any level of hierarchical structure but also at least one random effect above level one group (Bosker, R., & Snijders, T. A., 2011).It concerns a population with a hierarchical structure.

Multilevel modeling (also known as hierarchical linear modeling or mixed-effects modeling, nested data model, random effect, random parameter model) is a statistical technique used to analyze data with a hierarchical or nested structure. It accounts for the fact that data points are grouped or clustered within multiple levels, such as individuals within schools, patients within hospitals, or repeated measures (dash, 2023).The nesting, typically, but not always, is hierarchical. For instance, in this study, we have a two-level structure that would have many level-1(individual under-five children) units nested within a smaller number of level2 (cluster/EAs) units. The assumptions underlying the multilevel regression model are similar to the assumptions in ordinary multiple regression analysis: linear relationships, homoscedasticity, and normal distribution of the residuals. In multilevel analysis, strong assumptions are made: (i) the random effects are normal (or, if random slopes as well as random intercepts, that the joint distribution is multivariate normal), (ii) the model contains all relevant variables, (iii) there are enough observations at each level to really utilize the asymptotic theory results concerning the likelihood ratio test statistics and inverse of the information matrix as the estimator of the variances of the parameter estimates.

3.8.1. Multilevel Poisson Regression Model

Multilevel Poisson regression model is a nonlinear regression model often used for modelling count data. When the interest is to estimate intraclass correlation using a Poisson regression, a multilevel Poisson regression model that contains both fixed effects and random effects will be employed.

A multilevel count regression model can account for a lack of independence across levels of nested data (in our case, individual under-five children are nested within neighborhoods).

The multilevel count regression model has a count outcome (number of under-five children). Now consider the full model equation for the two-level Poisson regression for which individual under-five children are nested within EAs. The response variable, y_{ij} denotes the i^{th} individual within the j^{th} EA. Using a log link function, the two-level model is given by:

$$\text{Log}(\mu_{ij}) = \beta_0+\sum_{i=1}^{k} \beta_{lj}x_{lij} \quad i = 1,2,\ldots,k \qquad equ(3.2.1)$$

where $\beta_{0j} = \beta_0 + u_{0j}$, $\beta_{lj} = \beta_1 + u_{1j}+\ldots+\beta_{kj} + u_{kj}$

The two-level model can be rewritten as:

$\text{Log}(\mu_{ij}) = \beta_0 + \sum_{i=1}^{k} \beta_{l} x_{lij} + u_{0j} + \sum_{i=1}^{k} u_{ij} x_{lij}$ $\qquad equ(3.2.2)$

where x_{ij} $(x_{1ij}, x_{2ij,....} x_{kij})$ represent the first and the second level covariates, β_0=(β_0, β_1, β_2... β_k) are regression coefficients,$u_{0j,} u_{1j,} u_{2j}, \dots u_{kj}$ are the random effect of a model parameter at level two (communities level). It is assumed that the, $u_{0j,} u_{1j,} u_{2j}, \dots u_{kj}$ follow a normal distribution with mean zero and varianceσ_u^2. Without,$u_{0j,}$ $u_{1j} u_{2j}, \dots u_{kj}$, the equation can be considered as a single-level Poisson regression model.

3.8.2. Multilevel Negative Binomial Regression Model

Negative binomial regression is used for modeling count variables, usually over-dispersed count outcome variables. Count data with over-dispersion relative to a Poisson distribution are common in many biomedical applications. A popular approach to the analysis of such data is to use NB regression model. Often, because of the hierarchical study design or the data collection procedure, over dispersion and lack of independence may occur simultaneously, which makes the standard NB model adequate. To account for the over-dispersion and the inherent correlation of observations, a class of multilevel NB regression models with random effects is presented. The multilevel NB model is then generalized to cope with a more complex correlation structure. The multilevel NB model derives by allowing for regional random variations of the expected number of under-five children.

$\ln(\mu_{ij}) = \eta_{ij} + e_{ij}$ $\qquad equ(3.2.4)$

where $\text{cov}(e_{ij}, \eta_{ij}) = 0$ $\exp(e_{ij})$ follows a gamma probability distribution, $\Gamma(v)$, with mean 1 and variance $\alpha = v^{-1}$. The probability distribution

$P(Y_{ij}=y_{ij}) = \frac{\exp(-\exp(\eta\, ij + eij) \exp(\eta\, ij + eij)^{y_{ij}})}{y_{ij}} * v + y_{ij}$ $\qquad equ(3.2.5)$

One version of the multilevel negative binomial regression model is

$P(Y_{ij}=y_{ij}) = \frac{\Gamma(y_{ij}+v) V^{v} \mu_{ij}^{*y_{ij}}}{y_{ij}! \Gamma(v) (v+\mu_{ij})^{(v+\mu_{ij})}}$ $\qquad y_{ij} = 0,1,2.....$ $\qquad equ(3.2.6)$

With mean and variance given, respectively, as follows: The multilevel negative binomial regression model gives the expected mean of a number of under-five children.

$E(Y_{ij}) = \mu_{ij}^{*} = \log(\eta_{ij})$ its variance is given by $\text{Var}(Y_{ij}) = \mu_{ij} + \alpha\mu_{ij}^{2}$

Where $\eta_{ij} = \beta_{0j} + \beta_{1j} x_{1ij} + \beta_{2j} x_{2ij} + \cdots + \beta_{kj} x_{kij}$

V is the inverse dispersion parameter.

3.8.3 Multilevel ZIP Regression Model

The zero-inflated Poisson (ZIP) regression is used for count data that exhibit over dispersion and excess zeros. The data distribution combines the Poisson distribution and the logit distribution. The possible values of Y are the nonnegative integers: 0, 1, 2, 3, and so on.

Let y_{ij} be a count response for the i^{th} woman in the j^{th} community that follows a ZIP distribution. The multilevel ZIP model with random effect is given as follows

$$P(Y_{ij}=y_{ij}) = \begin{cases} \pi_{ij} + (1-\pi_{ij})e^{-\mu_{ij}}, & if\ y_{ij} = 0 \\ (1-\pi_{ij})\frac{e^{-\mu_{ij}}\mu_{ij}{}^{y_{ij}}}{y_{ij}\ !}, & if\ y_{ij} > 0 \end{cases} \qquad equ(3.2.7) \qquad 0 \le \pi_{ij} \le 1$$

Where $0 \le \pi_{ij} \le 1$ so that it incorporates more zeros than those permitted under the Poisson assumption (π_{ij} = 0), where as $\pi_{ij} < 0$ corresponds to the zero-deflated situation.

Recently, the ZIP regression model has been extended to the random effects setting, whereby random components and are incorporated within the logistic and Poisson linear predictors to account for the dependence of observations within j^{th} neighborhoods (Andy H. Lee, 2006)

These random effects ZIP models are region-specific in the sense that the random effects w_j and u_j so introduced are specific to the j^{th} neighborhood. In the following, a multi-level ZIP regression model is developed to handle correlated count data with extra zeros. Without loss of generality, consider the two-level hierarchical situation where Y_{ij} represents the i^{th} observation of a number of under-five child nested with the j^{th} neighborhoods (i=1, 2… n_i) , (j=1, 2…. m) and (k=1,2…..n_{ij}).

*L*et m be the total number of individuals in each neighborhood ($m=\sum_{i=1}^{n_i} n_{ij}$) and $n = \sum_{j=1}^{m}\sum_{i=1}^{n_i} n_{ij}$**,** gives the total number of observations. The observations may be taken to be independent between neighborhoods, but certainly, within-neighborhoods and within-individual correlations are anticipated, which can be modelled explicitly through random effects attached to the linear predictors:

$$\log(\mu_{ij}) = \beta_0 + \sum_{i=1}^{n_i} \beta_{kj} x_{kij} + u_{0j} + \sum_{i=1}^{n_i} U_{ij} x_{kij} \quad i = 1,2,\dots,n_i \qquad equ(3.2.8)$$

$$\text{logit}(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \gamma_0 + \sum_{i=1}^{n_i} \gamma_i z_{kij} + w_{0j} + \sum_{i=1}^{n_i} w_{ij} z_{kij} \qquad equ(3.2.9)$$

Here, the covariates x_{ij} and z_{ij} appearing in the respective Poisson and logistic components are not necessarily the same, β and γ are the corresponding vectors of regression coefficients (Andy H. Lee, 2006). For simplicity of presentation, the random effects u and w are assumed to be independent and normally distributed with mean zero and variance, δ_u^2 and δ_w^2 respectively.

3.8.4. Multilevel ZINB Regression Model

The zero-inflated negative binomial (ZINB) regression is used for count data that exhibit overdispersion and excess zeros.

A ZINB distribution arises as a mixture of a negative binomial and a distribution degenerated at zero and assigning a mass of π to extra zeros and a mass of (1-π) to a negative binomial distribution, where 0≤π≤1. Note that the negative binomial distribution is a continuous mixture of Poisson distributions, which allows the Poisson mean to be gamma distributed. (SB Javali, PV Pandit, 2010).

A multilevel ZINB regression incorporating random effects to account for data dependency and over-dispersion will be used (abbas Moghimbeigi & kazem Mohammad, 2008). Let Y_{ij} (i =1, 2, 3, n;j=1,2, 3, m) be the number of under-five children of the i^{th} household nested with j^{th} neighborhood and follows a ZINB distribution:

$$P(Y_{ij} = y_{ij}) = \begin{cases} \pi_{ij} + \frac{1-\pi_{ij}}{(1+\alpha\mu_{ij})^{\frac{1}{\alpha}}}, & \text{if } y_{ij} = 0 \\ 1 - \pi_{ij} \left(\frac{\Gamma\left(y_{ij}+\frac{1}{\alpha}\right)}{y_{ij}!\left(\frac{1}{\alpha}\right)}\left(1+\alpha\mu_{ij}\right)^{\frac{1}{\alpha}}\left(1+\frac{1}{\alpha\mu_{ij}}\right)^{-y_{ij}}\right., & \text{if } y_{ij} > 0 \end{cases} \qquad equ(3.2.10)$$

In this study, the number of under-five children in a household is nested in neighborhoods, and the number of under-five children is taken to be the response variable. *L*et n be the total number of individuals in each neighborhood ($n=\sum_{i=1}^{n_i} n_{ij}$) and $\sum_{j=1}^{m}\sum_{i=1}^{n_i} n_{ij}$ gives the total number of observations. Hence the responses of a number of under-five children who belong to different neighborhoods are independent, while they are correlated for those who live in the same neighborhoods. This dependence can be modelled explicitly by considering suitable random effects in the linear predictor. Negative binomial model for counts permit μ to depend on the explanatory variables. Then the two-level ZINB regression model can be expressed in vector form as:

$$\log(\mu_{ij}) = \beta_0 + \sum_{i=1}^{n_i} \beta_{kj} x_{kij} + u_{0j} + \sum_{i=1}^{n_i} U_{ij} x_{kij} \quad i = 1,2,\dots,n_i \qquad equ(3.2.11)$$

$$\text{logit}(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \gamma_0 + \sum_{i=1}^{n_i} \gamma_i z_{kij} + w_{0j} + \sum_{i=1}^{n_i} w_{ij} z_{kij} \qquad equ(3.2.12)$$

Here, the covariates x_{ij} and z_{ij} appearing in the respective Poisson and logistic components are not necessarily the same, β and γ are the corresponding vectors of regression coefficients (Andy H. Lee, 2006). For simplicity of presentation, the random effects u and w are assumed to be independent and normally distributed with mean zero and variance, σ_u^2 and σ_w^2 respectively.

3.8.5. Multilevel Hurdle Regression Model

A hurdle model is a class of statistical models where a random variable is modelled using two parts, the first which is the probability of attaining value 0, and the second part models the probability of the non-zero values. The uses of hurdle models are often motivated by an excess of zeroes in the data, that is not sufficiently accounted for in more standard statistical models.

Hurdle models were introduced by John G. Cragg in 1971, where the non-zero values of x were modelled using a normal model, and a probit model was used to model the zeros. The probit part of the model was said to model the presence of "hurdles" that must be overcome for the values of x to attain non-zero values, hence the designation hurdle model. Hurdle models were later developed for count data, with Poisson and negative binomial models for the non-zero counts.

A popular approach to the analysis of such data is to use a Hurdle Poisson regression model. Often, because of the hierarchical study design or the data collection procedure, zero-inflation and lack of independence may occur simultaneously, which the standard Hurdle Poisson regression model inadequate. To account for the preponderance of zero counts and the inherent correlation of observations, a class of multilevel Hurdle Poisson regression model with random effects is presented. Suppose that Yij is the number of under-five mortality for the i^{th} mother in the j^{th} region. Then, the multilevel Poisson Hurdle model can be written as ;

$$P(Y=y) = \begin{cases} \pi_{ij} & if\ y_{ij} = 0 \\ (1-\pi_{ij})\frac{\exp\left(-\mu_{ij})\mu_{ij}{}^{y_{ij}}\right)}{(1-\exp\left(-\mu_{ij}\right)y_{ij}!} & if\ y_{ij}>0 \end{cases} \qquad 0\leq\pi_{ij}\leq 1 \qquad equ(3.2.13)$$

Where $\Pi_{ij} = p(y_{ij}=0)$ and $\mu_{ij}=E(x_{ij}\beta)$

3.8.7 Parameter Estimation

Once a model is specified with its parameters, and data have been collected, one is in a position to evaluate its goodness of fit, that is, how well it fits the observed data. Goodness of fit is assessed by finding parameter values of a model that best fits the data, a procedure called parameter estimation. There are two general methods of parameter estimation. They are least-squares estimation (LSE) and maximum likelihood estimation (MLE). MLE is a standard approach to parameter estimation and inference in statistics. MLE has many optimal properties in estimation: sufficiency (complete information about the parameter of interest contained in its MLE estimator); consistency (true parameter value that generated the data recovered asymptotically, i.e. for data of sufficiently large samples); efficiency (lowest-possible variance of parameter estimates achieved asymptotically); and parameterization invariance (same MLE solution obtained independent of the parameterization used). Generally, generalized linear models (GLMs) do not have a closed-form solution for maximizing the likelihood function. Therefore, to approximate the maximum likelihood estimates (MLEs) of GLMs, we often rely on the Newton-Raphson algorithm. The log-likelihood function for the coefficient β is expressed as follows:

$$l(\beta,\phi) = \sum\left\{\frac{y_i\theta_i - b(\theta_i)}{\phi} + c(y_i, \phi)\right\} = \sum l_i(\theta_i, \phi) = \sum l_i \qquad equ(3.2.15)$$

where $\theta_i = \theta(x', \beta) = \theta(\eta_i)$

The MLE of β are obtained by maximizing the log-likelihood functions l(β,ϕ) .

where θ is known,

The likelihood function of GLM depends on β only through the linear predictor η_i.The MLE of β are the solutions of the simultaneous equations of

$$\frac{\partial l(\beta,\phi)}{\partial \beta} = \sum \frac{\partial l_i}{\partial \beta_i} = 0$$

(Alexandre Brouste, 2020)

3.8.8 The ICC

Intra class correlation coefficient (ICC) (also termed variance partition coefficient (VPC) in its most general form) is relevant for understanding contextual phenomena expressed with continuous variables. In the linear case, the ICC informs us on the proportion of total variance in the outcome that is attributable to the area level.

In the linear model, the ICC is based on the clear distinction that exists between the individual level variance and the area level variance. Indeed, knowing the mean value of a continuous outcome variable in each area, would not allow to infer the values of the variable for each individual: The individual level variance within areas could be small or very large.

$$\text{ICC} = \frac{V_A}{(V_A + V_I)} \qquad equ(3.2.3)$$

where V_I is the variance in the number of child death for women of the reproductive age (first level variance), and V_A is the variance in Under-five child mortality among clusters/ EAs (second level variance). If the Intraclass Correlation Coefficient (ICC) is equal to 1, it suggests that there is a perfect correlation or similarity in Under-five child mortality (U5CM) between two children occupying the same cluster or Enumeration Area (EA). In other words, when the ICC is 1, there is no variability or difference in U5CM within a cluster. All children within the same cluster have identical U5CM rates. This indicates a high degree of similarity or homogeneity in terms of child mortality within the clusters. (that is, 100 % of the total death of child differences are at the cluster level), and an ICC equal to 0 would show that the cluster does not share any U5CM related common level of Childs.

When the ICC is 0, the suitability of performing a multilevel analysis is less obvious. In the absence of a multilevel structure, a single-level individual analysis is appropriate. We shall use adjusted intraclass correlation (ICC_{Adj}) to describe the autocorrelation after taking into account cluster level variables. This measure is of relevance, as it quantifies the autocorrelation of individual rates of death of children within clusters.

The ICC_{Adj} is the proportion of total variance in the rate of death of children that remains at the cluster level after taking into account the cluster level composition of the number of death of Children.

3.8.9 Goodness of Fit Tests

A goodness-of-fit test**,** in general, refers to measuring how well do the observed data correspond to the fitted (assumed) model. Like in linear regression, in essence, the goodness-of-fit test compares the observed values to the expected (fitted or predicted) values.

A goodness-of-fit statistic tests the following hypothesis:

H_0: the model fits vs. H_1: the model does not fit

3.8.9.1 Wald test

The Wald test (also called the Wald Chi-Squared Test) is a way to find out if explanatory variables are significant. The Wald statistic can be used to assess the contribution of individual predictors or the significance of individual coefficients in a given model. The Wald statistic is the ratio of the square of the estimated regression coefficient to the square of the standard error of the coefficient. The Wald statistic is asymptotically distributed as a chi-square distribution and each Wald statistic is compared with a chi-square with 1 degree of freedom.

$$w^2 = \left(\frac{\hat{\beta}_l}{SE(\hat{\beta}_l)}\right)^2 = \left(\frac{\hat{\beta}_l}{\sqrt{var(\hat{\beta}_l)}}\right)^2 \sim x^2 \qquad equ(3.2.17)$$

100(1-α)% confidence interval for β_i is given by

$$\hat{\beta}_i \pm z_{\frac{\alpha}{2}} \sqrt{var(\hat{\beta}_l)}$$

3.8.10 Model Comparisons

In order to select the best fitting model, Akakie Information criteria (AIC) and Bayesian information criterion (BIC) have been used.

3.8.10.1 Akakie Information criteria (AIC)

AIC is used to compare different possible models and determine the best fitting model for the data.

The AIC is computed as:

$$AIC = -2\ln L + 2p \qquad equ(3.2.18)$$

where p is the number of unknown parameters included in the model and lnL denotes the log likelihood of a model that is to be compared with the other models.

A preferred model is characterized by having the lowest AIC value.

3.8.10.2 Bayesian information criterion (BIC).

The BIC was developed by Gideon E. Schwarz, who gave a Bayesian argument for adopting it. The Bayesian Information Criterion (BIC) is a measure of model fit that takes into account goodness of fit and the number of parameters used to achieve the fit. The BIC is a well-known general approach to model selection that favors more parsimonious models over more complex models (i.e., it adds a penalty based on the number of parameters being estimated in the model). Unlike the Akakie information criteria, the Bayesian information criteria (BIC) takes into account the size of the data under consideration. One form for calculating the BIC is given by (Bauldry, 2015)

$$BIC = -2\ln L + k\log(n) \qquad equ(3.2.19)$$

where n is the sample size of the data, lnL denotes the log likelihood of a model that is to be compared with the other model , k is the number of parameters in the model including the intercept and n is sample size of the data. The BIC can be used to assess two competing models.

3.8.11 Test for Comparison of Non-nested Models

3.8.11.1 Vuong Test

The Vuong non-nested test is based on a comparison of the predicted probabilities of two models that do not nest. Examples include comparisons of zero-inflated count models with their non-zero inflated analogs (e.g., zero-inflated Poisson versus ordinary Poisson, or zero-inflated negative-binomial versus ordinary negative-binomial). The Vuong test for non-nested models is being widely used as a test of zero-inflation.

Let

$$M_i=\log(p_1(y_i/x_i)/p_2(y_i/x_i))$$

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, $P(y_i/x_i)$ are the predicted probabilities of observed count for case from model, then the Vuong test statistic is simply the average log-likelihood ratio suitably normalized.

The Vuong test statistics is

$$V = \sqrt{n}\frac{\frac{\sum_{i=1}^{n} m_i}{n}}{\sqrt{\sum_{i=1}^{n}(m_i - m)^2}} \qquad equ(3.2.20)$$

where m is the mean of m_i and n sample size. The hypotheses of the Vuong test are:

H_0=E(m_i) = 0 Vs H_1=(m_i) ± 0

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

If V>Zα/2, the first model is preferred.

If V< −Zα /2, the second model is preferred.

If |V|<Zα/2, none of the models are preferred.

CHAPTER FOUR

RESULT AND DISCUSSIONS

4.1 Descriptive Statistics

Information on the number of deaths of under-five children obtained from a total of 5753 women in Ethiopia was studied. Table 1 shows the frequency and percentage distribution of the number of under-5 mortality in Ethiopia. While 75.54% of the women never experienced under-5 death of their children, 15.75 %, 6%, 1.91%, 0.42%, 0.1%, 0.21%, and 0.07% of them lost 1, 2, 3, 4, 5, 6, and 7 of their under- five children, respectively. The mean value represents the average number of under-five mortality per mother in Ethiopia. The mean 0.37, on average mothers in Ethiopia experience less than one under-five mortality per mother. A variance of 0.64 indicates that there is relatively moderate level of variability in the number of under-five mortality per mother in Ethiopia. Since the number of zero outcomes is large, the histogram is highly picked at the very beginning (about the zero values) (Figure1). However, larger numbers of under-five deaths per mother are less frequently observed showing a positively (or right) skewed distribution. Further screening of the number of under-five death calculated showed that the variance (0.64) is greater than the mean (0.37 indicating over-dispersion. This is an indication that the data could be fitted better by count data models which takes into account excess zeroes (table 1 and figure 1).

Table 4.1: Frequency and percentage distribution of the number of under-5 mortality per mother in Ethiopia

Under-five child death per mother	Frequency	Percent
0	4346	75.54
1	906	15.75
2	345	6
3	110	1.91
4	24	0.42
5	6	0.1
6	12	0.21
7	4	0.07
Total	5753	100
Mean	0.37	
Variance	0.64	

Figure 4.1: Histogram of percentage of under-5 deaths per mother

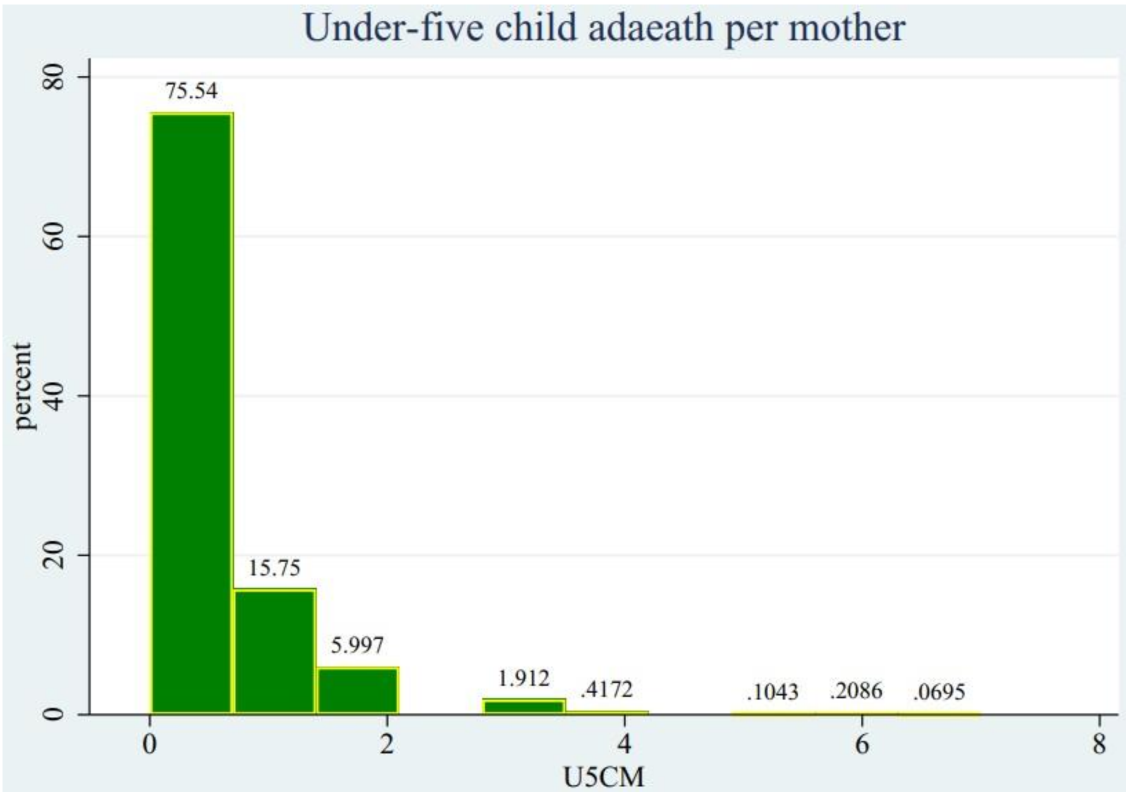


Table 4.2 below presents further detailed statistics of under-five death by category for each covariate.

From a total of mothers 5753, 2143 (37.3%), 1915 (33.3), 1156(20%), and 539(9.4%) mothers birth their children in spring, winter, autumn and summer season respectively. And 22.4 %, 26.2%, 24,8% and 27.3% death occurs during a spring, winter, autumn and summer season respectively. These statistics provide information about the distribution of death occurrences based on the season of birth. It appears that the highest death rate is observed in the winter season (26.2%), while the lowest death rate is observed in the summer season (27.3%). Based on summary of the region, it allows for a comparison of death rates across different regions. The region with the highest death rate is Benishangul Gumuz (34.3%), while the region with the lowest death rate is Tigray (15.4%).

Regarding educational level, It appears that children whose mothers were not educated have a higher proportion of deaths (30%) compared to those with different educational levels. When we talk about distance to get water water accessible with in 30 minute have higher child death rate 26.8% and water accessible on premise have lower death rate 14.2%.

From a total of respondents, 1607(27.9%) of mothers have 1-4 Number of house member, 2691 (46.8%) of mothers have 5-7 household members and 1455(25.3%) of mothers have above 7 house hold member. The highest death rate 697(25.9%) occurs mothers who have 5-7 household member and lower death rate 293(18.2%) mothers who have 1-4 household member. From the total of respondents, 4598(79.9%) are male household headed and 1155(20.1%) are female household headed. Male household headed have 1155(25.1%) death rate and 252 (21.8%) is death rate for female household headed.

Younger age of mothers at first birth have 749(30.2) higher child death rate than that of mothers older age at first birth 156 (24.3%). From a total of mothers 3874 (67.3%) were not used contraceptive methods and 1879(32.7%) were used contraceptive methods. The highest death 1038(26.8%) were occurs mothers who have not used contraceptive method and lower death rate 369(19.9%) occurs mothers who have used contraceptive methods. Married mothers have 1319(24.6%) death of child occurrence and not married mothers have 88(22.1%) have child death. Birth order increase number of child death also increase, the first birth order have 104(8.1%) death and 668(51.9%) death occurs child born on the 6th and above birth orders.

Regarding to types of birth, single birth have 1328(23.8%) of child death and multiple birth have 74(44.3%) death occurrence.

Table 4.2 Summary statistics of all covariates that are included in the analysis

covariates	Category	No death n(%)	Death n (%)	Total n
Season of birth	Spring	1663 (77.6)	480 (22.4)	2143
	Winter	1413 (73.8)	502(26.2)	1915
	Autumn	878 (75.2)	278(24.8)	1156
	Summer	392 (72.7)	147(27.3)	539
Region	Tigray	384 (84.6)	70(15.4)	454
	Afar	521 (79.9)	131(20.1)	652
	Amhara	390 (76.3)	121(23.7)	511
	Oromia	513(71.3)	206(28.7)	719
	Somali	428 (67.2)	209(32.8)	637
	Benishangul gumuz	348 (65.7)	182(34.3)	530
	SNNP	508 (77)	152(23)	660
	Gambela	322 (71.6)	128(28.4)	450
	Hareri	341 (76.3)	106(23.7)	447
	Addis Ababa	276 (94.8)	15(5.2)	291
	Dire dawa	315(78.4)	87(21.6)	402
Highest educational Level of mother	Not educated	2203(70)	946(30)	3149
	Elementary	1444(79.2)	379(20.8)	1823
	Secondary	432(90)	48(10)	480
	Higher	267(88.7)	34(11.3)	301
distance to get water	<=30 minute	2089(73.2)	766(26.8)	2855
	Above 30 minute	1323(73.1)	486(26.9)	1809
	On premises	934(85.8)	155(14.2)	1089

HH size	1-4	1314(81.7)	293(18.2)	1607
	5-7	1994(74.1)	697(25.9)	2691
	Above 7	1038(71.3)	417(28.7)	1455
Sex household head	Male	3443(74.9)	1155(25.1)	4598
	Female	903(78.2)	252(21.8)	1155
Age household head	<35	2477(79.6)	634(20.4)	3111
	35 to 50	1382(69.1)	617(30.9)	1999
	Above 50	487(75.7)	156(24.3)	643
age of mothers at 1st birth	<19	1730(69.8)	749(30.2)	2479
	19 to 35	2599(79.9)	654(20.1)	3253
	35 and above	17(81)	4(19)	21
Current Contraceptive use	not used	2836(73.2)	1038(26.8)	3874
	Used	1510(80.4)	369(19.6)	1879
Marital status of mother	Not married	310(77.9)	88(22.1)	398
	Married	4036(75.4)	1319(24.6)	5355
Birth order	First	1157(91.8)	104(8.2)	1261
	2nd and 3rd	1630(85.7)	272(14.3)	1902
	4th and 5th	940(72.1)	363(27.9)	1303
	6 and above	619(48.1)	668(51.9)	1287
Type of birth	Single	4258(76.2)	1328(23.8)	5586
	Multiple	88(55.7)	74(44.3)	167
Sex child	Male	2230(75.1)	739(24.9)	2969
	Female	2116(76)	668(23.9)	2784
Place of delivery	Home	2093(70.8)	862(29.2)	2955
	At health facility	2253(80.5)	545(19.5)	2798
Birth interval	< 2 year	1465(80.4)	357(19.6)	1822
	2 up to 4	1519(72.9)	565(27.1)	2084
	Above 4 year	1362(73.7)	485(26.3)	1847

The total of 5753 women included, 2955 (51.4%) have delivered at home and 2798 (48.6%) at any health center. mothers delivered at home have about 862(29.2%) experienced death of a child and mothers who have delivered at health facility have 545(19.6%) experience of child death. Mothers delivered at home have higher under-five child death experience than that of mothers delivered at health facility. From a total. Mothers with a birth interval of less than 2 years have a higher proportion of under-five deaths (19.6%) compared to those with longer birth intervals.

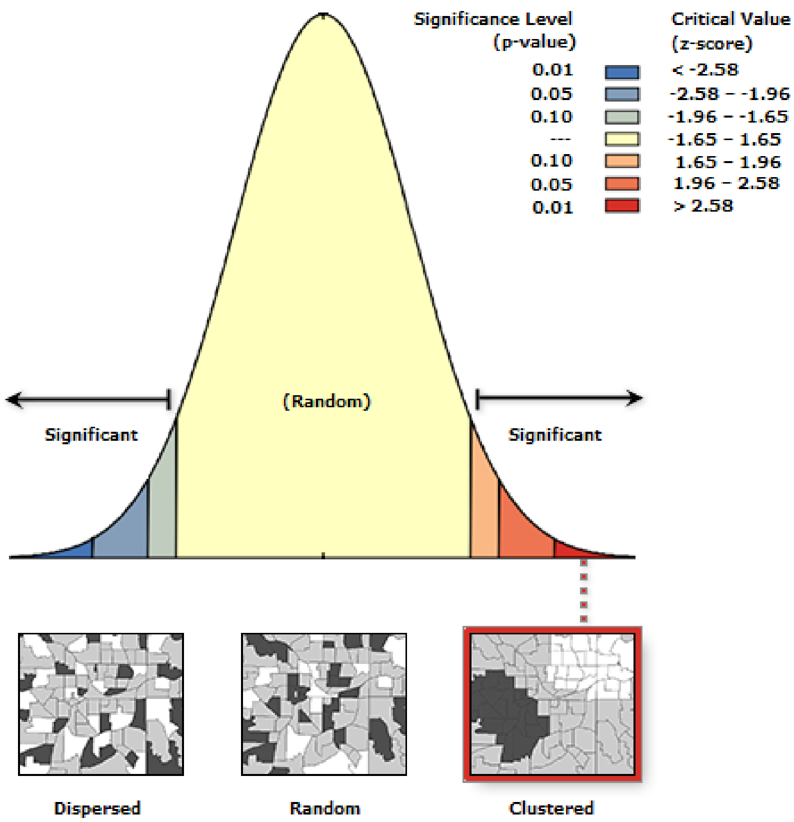
4.2 Spatial Analysis

4.2.1 Global Measures of Spatial Autocorrelation

In Ethiopia, spatial patterns were discovered. The worldwide Moran's I value (I = 0.416267, z score =5.410983, P value=0.0000 < 0.0001) revealed a substantial group of under-five mortality of child throughout the country. Under-five death had a statistically significant regional clustering pattern among Ethiopian children and there is less than 0.1% likelihood that this clustered pattern could be the result of random chance.

Figure 4.2: Spatial autocorrelation based on feature locations and attribute values using the Global Moran I statistic

Moran's Index: 0.245055
z-score: 5.410983
p-value: 0.000000

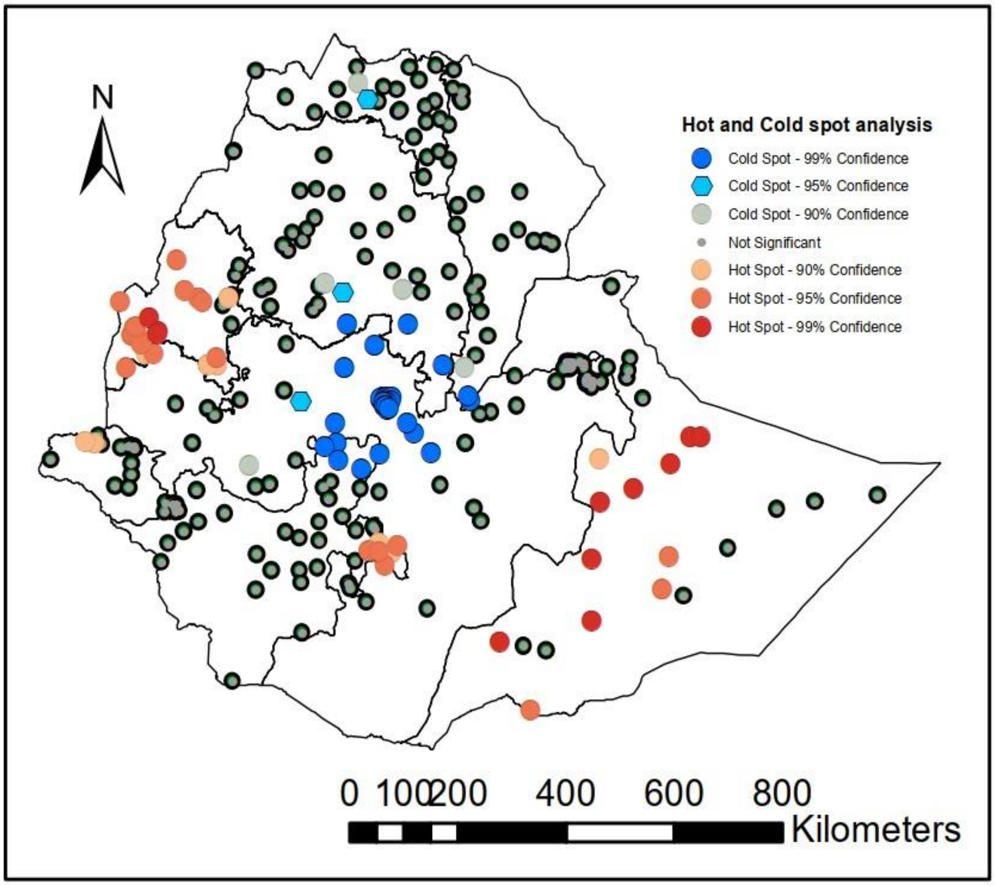


4.3 Local Indicators of Spatial Autocorrelation

4.3.1 Hot and Cold Spot Analysis

Hotspot and coldSpot analysis was performed. Benshangul Gumez, Somali and eastern part of SNNP were identified as statistically significant hot spot regions (risky areas of the study), whereas Addis Ababa, central oromia south Amhara, northern Tigre and south Afar were identified as statistically significant cold spot regions (non-risky areas of the study) (Figure 3).

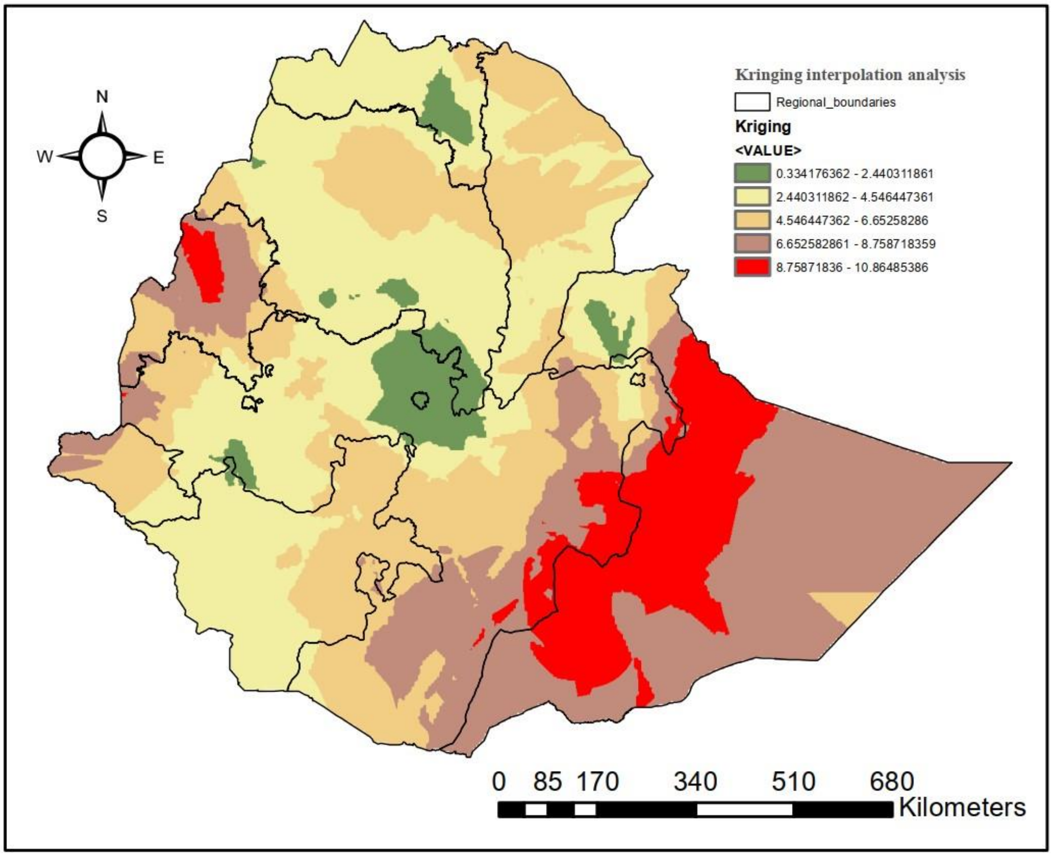
Figure 4. 3: Hotspot and cold spot analysis of under-five child mortality in Ethiopia, EMDHS 2019.



4.3.2 Spatial Interpolation

Based on figure 4, the spatial kriging interpolation analysis in figure 4 in next page predicted high-risk regions for under-five child death. The prediction of high-risk areas was indicated by red predictions. North west Somali, Benishangul gumuz, west Gambela, western Oromia are predicted to be more risky regions compared to other regions. Conversely, south west Amhara, Nothern peak of Afar, central Amhara, south and central Oromia and Addis Abeba were identified as less risk areas of the country.

Figure 4.4: Empirical Bayesian kriging interpolation of under-five death in Ethiopia, EMDHS 2019.



4.4 Multilevel Count Analysis of the Data

In our multilevel analysis, we employed a two-level structure since the dataset exhibits a hierarchical structure where 5753 mothers at level 1 are nested within 305 enumeration areas at level 2.

4.4.1 Intra Class Correlation

The intra-class correlation (ICC) coefficient has been estimated to be 0.111 indicating that a greater share of the total variation in the number of under-five mortality per mother is associated with cluster membership.

4.4.2 Goodness of Fit and Model Selection Criteria

Six multilevel count regression models were fitted with the results shown in Table 3. As can be seen in the table, the MZINB model has the least AIC, BIC and deviance values'. As a result, we can confidently assert that, in this study, the multilevel ZINB regression model outperforms the other models (the MZINB model offers a significantly better fit to the data.

Table 4.3: Goodness of fit and model selection criteria

Model	AIC	BIC	Deviance
MP	9274.6	9287.93	9270
MNB	8948.18	8968.16	8942.18
MZIP	9109.87	9129.85	9103.87
MZINB	8919.96	8946.59	8911.96
MHP	9232.44	9252.4	9226.44
MHNB	9111.35	9137.98	9103.35

4.4.3 Vuong Test

The first comparison is made between the Poisson model and the NB model, with a Vuong test statistic of 3.649 with p-value <0.0003 implying that the NB model is preferred to the Poisson model for predicting the number of under-five mortality of child per mother.
The Vuong statistic for the ZIP versus Poisson (3.276, p-value <001) favors the ZIP model. When we compare ZIP versus NB with (voung statistics = -2.107, p-value = 0.01<0.05) and model (NB) is favored. After a series of tests and model comparisons (as shown in Table 4) the ZINB modelis best model to fit the number of under-five mortality of child per mother.

The Vuong statistic for the ZINB versus NB (5.206, p-value <0.00000) favors the ZINB model. Based on the above vuong non nested test statistics results in table 4 ZINB modes is best model to fit the number of under-five mortality of child per mother than ZIP, NB, HP, HNB, POISSON models. This idea also supported by deviance, AIC AND BIC goodness of fit test. Then ZINB is preferable model for fitting under-five child mortality per mother.

Table 4.4: Vuong non- nested tests results

Model Comparison	Vuong Test Statistic	P-value	Preferable Mode
NB vs poisson	3.649	0.00013151	NB
ZIP vs poisson	3.276	0.00052642	ZIP
Zip vs NB	-2.107	0.017547	NB
ZINB vs NB	5.209	0.00000	ZINB
ZIP VS HP	-4.033	0.00000	HP
ZINB vs ZIP	1.787	0.03697	ZINB
ZINB vs HP	1.775	0.037918	ZINB
ZINB vs HNB	3.417	0.00031616	ZINB

In this study, various software have been used for the analysis of the data. For the mapping and analysis of spatial data, Geographic Information System (GIS), specifically ArcGIS version 10.8 has been used. This software provided the necessary functionalities to work with spatial data, allowing visualizing and analyzing geographic information.

To fit multilevel zero-inflated Negative binomial (ZINB) model, R version 4.3.3 and STATA version 16 have been used.

By leveraging the capabilities of Geographic Information System (ArcGIS) as well as R and STATA statistical software it was possible to conduct a comprehensive analysis of our data.

4.6 Results of Multilevel ZINB Regression

The selected MZINB regression model fit results generated two separate models: the first one corresponds to the count part and the second corresponds to the zero-inflated part.

4.6.1 **The Count Part**

The expected number of under-five child mortality in female headed households is 1.27 times the expected number of under-five deaths in male headed households (IRR=1.27) . That is, the expected number of under-five deaths in female headed households increased by 27% as compared to that in male headed households) holding all other variables in the model constant. Similarly, the expected number of under-five death decreased by a factor of 0.6088 (IRR = 0.6088) and 0.4708in households of size 5-7 and above 7 members respectively, as compared to the expected number in households of size 1-4 members.

The estimated IRR corresponding to household heads aged 35-50 and above 50 are 1.241 and 1.1571 respectively. This means that, on average, the expected number of under-five deaths for 35-50 and above 50 years of age household heads are 1.241 and 1.1571 times the expected number of under-five deaths in households with heads under 35 years of age.

Compared to the reference category (available in premises), when the time to get water is less than or equal to 30 minutes, there is an estimated IRR of 1.2616. This means that, on average, the expected number of under-five child death per mother when it takes at most minutes to get water is 1.2616 times the expected number of under-five deaths when the water is available on premises. Likewise, when the time to get water takes above 30 minutes, there is an estimated IRR of 0.5779. This means that, the expected number of under-five deaths when it takes more than 30 minutes to get water is 0.5779 times the expected number of under-five deaths when the water is available on premises. Considering contraceptive use, the expected number of under-five deaths when the mother uses contraceptive methods decreased by a factor of 0.79 (IRR=0.79) compared to the expected number of under-five deaths when the mother is not using contraception

Mothers who had their first child at age 35 and above have an estimated IRR of 8.6279. This indicates that, the expected number of under-five deaths for mothers who had their first child at age 35 and above is 8.6279 times the expected number of under-five deaths for mother who had their first child when they are younger than 19 years. Compared to the reference category (first birth), a birth order of 2nd and 3rd, has an estimated IRR of 2.1012 implying that the expected number of under-five child death per mother for a birth order of 2nd and 3rd is 2.1012 times that for a first birth. Also, the expected number of under-five deaths for a birth order of 4th and 5^{th} is 4.0812 times the expected number of under-five deaths for a first birth. Moreover, multiple birth increases the incidence rate of under-five mortality (IRR=1.5595).

Table 4.5: Parameter estimates of MZINB regression model

Variables	Estimate	IRR	S. E	z –value	P-value	95% CI	
						Lower	Upper
Intercept	-1.6390	0.1942	0.4717	-3.475	0.000511	-2.5634	-0.7146
Season of birth(ref. spring)							
Winter	0.0145	1.0146	0.0913	0.159	0.87364	-0.1644	0.1934
Autumn	0.0351	1.0357	0.1106	0.318	0.75074	-0.1817	0.2520
Summer	0.1306	1.1395	0.1399	0.933	0.35079	-0.1437	0.4049
Region (Tigray)							
Afar	0.0587	1.0604	0.2867	0.205	0.83793	-0.5038	0.6212
Amhara	0.1653	1.1797	0.2813	0.588	0.55674	-0.3860	0.7165
Oromiya	0.0265	1.0268	0.2695	0.098	0.92157	-0.5017	0.5548
Somali	0.0520	1.0534	0.2725	0.191	0.84883	-0.4822	0.5861
Benishangul gumuz	0.3275	1.3875	0.2724	1.202	0.22927	-0.2064	0.8615
SNNP	0.0928	1.0973	0.2743	0.338	0.73517	-0.4448	0.6304
Gambela	0.2448	1.2774	0.2849	0.859	0.39024	-0.3136	0.8032
Hareri	0.0035	1.004	0.3021	0.012	0.99067	-0.5885	0.5956
Addis Ababa	0.1189	1.1263	0.5823	0.204	0.83842	-1.0237	1.2614
Dire dawa	0.2079	1.2310	0.2998	0.694	0.48795	-0.3796	0.7954
Highest edu. Level of mother (ref. higher)							
Not educated	-0.0243	0.9759	0.1032	-0.235	0.81382	-0.2266	0.1780
Elementary	-0.5799	0.5599	0.3509	-1.653	0.09837	-1.2675	0.1078
Secondary	-0.9075	0.4035	0.5960	-1.523	0.12780	-2.0756	0.2605
Source of drinking water(ref. unimproved)							
Improved	0.0717	1.0743	0.0953	0.752	0.452008	-0.1152	0.2586
distance to get water(ref. on premises)							
<=30 minute	0.2324	1.2616	0.0939	2.475	0.01331	0.0484	0.4164
Above 30 minute	-0.5483	0.5779	0.2177	-2.519	0.01177	-0.9750	-0.1216
Type toilet facility(Ref. not improved)							
No toilet	0.2631	1.301	0.1428	1.842	0.065486	-0.0169	0.5430
Improved toilet	0.0099	1.0096	0.1036	0.097	0.923105	-0.1930	0.2130

HH size(ref. 1-4)							
5-7	-0.4962	0.6088	0.1427	-3.476	0.00050	-0.7759	-0.2164
Above 7	-0.7532	0.4708	0.1569	-4.802	0.000001	-1.0607	-0.4458
Sex household head (ref. male)							
Female	0.23927	1.2703	0.1173	2.040	0.04133	0.0094	0.4691
Age household head(ref. <35)							
35 to 50	0.21593	1.241	0.0977	2.210	0.027117	0.0244	0.4074
Above 50	0.1459	1.1571	0.1388	1.051	0.293262	-0.1262	0.4179
age of mothers at 1st birth (ref. <19)							
19 to 35	-0.0760	0.9268	0.0820	-0.926	0.354369	-0.2367	0.0848
35 and above	2.1550	8.6279	0.5239	4.113	0.000039	1.1281	3.1818
Current Contraceptive use(ref. not used)							
Used	-0.2343	0.7911	0.0992	-2.361	0.018222	-0.4288	-0.0398
Marital status of mother (ref. not married)							
Married	0.2250	1.2523	0.1842	1.222	0.221875	-0.1360	0.5860
Birth order (ref. first birth)							
2^{nd} and 3^{rd}	0.7425	2.1012	0.3332	2.229	0.025835	0.0895	1.3955
4^{th} and 5^{th}	1.4064	4.0812	0.3369	4.174	0.000029	0.7461	2.0667
6 and above	2.2755	9.728	0.3378	6.736	0.00000	1.6134	2.9376
Type of birth (ref. single)							
multiple	0.4444	1.5595	0.1334	3.331	0.00086	0.1829	0.7059
Sex child (ref. male)							
Female	0.0423	1.0432	0.0735	0.576	0.564916	-0.1018	0.1865
Place of delivery(ref. at home)							
At health facility	-0.0949	0.9095	0.0898	-1.057	0.29039	-0.2708	0.0810
Birth interval(ref.< 2 year)							
2 up to 4	-0.1396	0.8697	0.0932	-1.498	0.13410	-0.3221	0.0430
Above 4 year	-0.1418	0.8678	0.0989	-1.433	0.15184	-0.3356	0.0521
Zero inflated part							
Intercept	2.10519	8.2145	0.2924	7.199	0.000001	1.5321	2.6783
Season of birth(ref spring)							
Winter	-0.2188	0.8034	0.0831	-2.634	0.008436	-0.3817	-0.0560
Autumn	-0.13276	0.8757	0.0972	-1.366	0.172088	-0.3233	0.0578
Summer	-0.51165	0.5995	0.1272	-4.023	<0.00001	-0.7609	-0.2624
Region(Tigray)							
Afar	-0.31739	0.7280	0.1866	1.701	0.088976	-0.6831	0.0484
Amara	-0.45425	0.6349	0.1876	-2.421	0.015462	-0.8219	-0.0866

Oromia	-0.72992	0.4819	0.1754	-4.163	0.000001	-1.0736	-0.3862
Somali	-0.81114	0.4444	0.1858	-4.367	0.0000	-1.1752	-0.4471
Benishangul Gumuz	-1.04534	0.3516	0.1844	-5.670	0.00000	-1.4067	-0.6840
SNNP	-0.4164	0.6594	0.1836	-2.268	0.02332	-0.7763	-0.0566
Gambela	-0.9013	0.4060	0.1929	-4.673	0.00000	-1.2793	-0.5232
Hareri	-0.7523	0.4713	0.1937	-3.884	0.000103	-1.1319	-0.3727
Addis ababa	0.42484	1.5293	0.3228	1.316	0.188149	-0.2079	1.0575
Dire dawa	-0.62204	0.5368	0.2019	-3.081	0.002064	-1.0178	-0.2263
Highest edu. Level of mother (ref. higher)							
Not educated	-0.09907	0.9057	0.0885	-1.120	0.262791	-0.2725	0.0743
Lower	0.25898	1.2956	0.1831	1.414	0.157282	-0.0999	0.6179
Secondary	-0.15590	0.8556	0.2215	-0.704	0.481433	-0.5899	0.2781
Source of drinking water(ref. unimproved)							
Improved	-0.07736	0.9256	0.0792	-0.977	0.328657	-0.2326	0.0779
Time to get water(ref. on premises)							
<=30 minute	-0.06386	0.9381	0.0866	-0.737	0.461058	-0.2337	0.1059
Above 30 minute	0.10205	1.1074	0.1386	0.736	0.461555	-0.1696	0.3737
Type toilet facility(Ref. not improved)							
No toilet	0.09602	1.1008	0.1242	0.773	0.439394	-0.1474	0.3394
Improved toilet	0.17799	1.1948	0.0882	2.018	0.043573	0.0051	0.3508
HH size(ref. 1-4)							
5-7	0.92784	2.5290	0.1114	8.333	0.00000	0.7096	1.1461
Above 7	2.16824	8.7423	0.1479	14.659	0.00000	1.8783	2.4581
Sex of household head (ref. male)							
Female	0.32316	1.3815	0.1066	3.032	0.002431	0.1142	0.5321
Age household head(ref. <35)							
35 to 50	-0.03311	0.9674	0.0858	0.386	0.699595	-0.2013	0.1351
Above 50	-0.11661	0.8899	0.1271	-0.918	0.358692	-0.3656	0.1324
age of mothers at 1st birth (ref. <19)							
19 to 35	0.07380	1.0766	0.0738	1.358	0.174574	-0.0445	0.2449
35 and above	-0.51813	0.5956	0.6149	-0.843	0.399485	-1.7234	0.6872
Current Contraceptive use(ref. not used)							
Used	0.11609	1.1231	0.0842	1.380	0.167739	-0.0488	0.2810
Marital status of mother (ref. not married)							
Married	0.24678	1.2799	0.1642	1.503	0.132791	-0.0750	0.5685
Birth order (ref. first birth)							
2^{nd} and 3^{rd}	-0.98341	0.3738	0.1413	-6.961	0.00000	-1.2603	-0.7065

4^{th} and 5^{th}	-2.21494	0.1092	0.1625	-13.633	0.00000	-2.5334	-1.8965
6 and above	-3.89354	0.0204	0.1808	-21.532	0.00000	-4.2480	-3.5391
Type of birth (ref. single)							
multiple	-0.88337	0.4134	0.1833	-4.819	0.00000	-1.2427	-0.5241
Sex child (ref. male)							
Female	0.07985	1.0831	0.0696	1.147	0.251386	-0.0566	0.2163
Place of delivery(ref. at home)							
At health facility	0.20142	1.2231	0.0826	2.439	0.014711	0.0396	0.3633
Birth interval(ref.< 2 year)							
2 up to 4	0.24403	1.2764	0.0960	2.542	0.011018	0.0559	0.4322
Above 4 year	0.18580	1.2041	0.1029	1.805	0.071052	-0.0159	0.3875

4.6.2 Zero Inflated Part

Table 4.5 (zero inflated part) also presents the estimated zero-inflated negative binomial (ZINB) regression model fit results for selected covariates, specifically focusing on the "always 0 group." The coefficients in this group represent the factor change in the odds of belonging to the always-0 group compared to the not always-0 group.

Based on the results in table 4, the estimated odds that the number of zero under-five death for winter and summer season are 0.8034 and 0.5995 times lower (IRR=0.8034, 95% CI:-0.3812 to -0.0578 and IRR=0.5995, 95% CI: -0.761 to -0.262) to the spring season respectively. The probability of observing excessive probability of under-five death becomes zero in Addis Ababa is 1.5293 times as compared to Tigray. Similarly, the estimated odds that the number of under-five death becomes zero for mothers who located in Amhara, Oromia, Somali, Benishangul gumuz, SNNP , Gambela, Hareri, and Dire dawa is 0.6349, 0.4819, 0.4444, 0.3516, 0.6594, 0.4060, 0.4713 and 0.5368 times for women in Tigray.

In terms of the log odds, household sizes of 5-7 and above 7 members are positively associated with excess zeros indicating a higher likelihood of nonexistence of under-five child mortality, with an estimated log odds of 0.92784 ($p < 0.05$) and 0.16824 ($p < 0.05$) respectively.

Considering “birth order of child”, The estimated odds that the number of U5 death becomes zero for the second and third-born, fourth and fifth-born and sixth-born and above decreased by 63% (IRR=0.37 CI: -1.26 up to -0.71), 89% (IRR=0.11 CI: -2.53 up to -1.89) 98% (IRR=0.02 CI: -4.25 up to -3.54) respectively as compared to first-born children (the reference category). As the birth order increases, the probability of under-five death of non-zero counts for child mortality decreases significantly. Multiple births are significantly associated with a lower probability of being in the excess zero group for child mortality compared to single births with log odds of -0.88337 (p = 0.000). The 95% confidence interval for the estimate (-1.2427 to -0.5241) further confirms the significance of the association. Compared to deliveries at home (the reference category), deliveries at a health facility estimated odds that the number of U5 death becomes zero is increased by 22.3% (IRR=1.223, 95% CI: 0.039 – 0.363). Indicating a positive statistically significant association with excess zeros. Understanding this association can inform efforts to improve the quality of care during facility-based deliveries and reduce child mortality rates.

Birth intervals of 2 up to 4 years are associated with a higher probability of non-zero counts for child mortality compared to birth intervals of less than 2 years with log odds of 0.18580 (p = 0.011018), although the 95% confidence interval for the estimate (-0.0159 to 0.3875) suggests a potential positive impact. Birth intervals above 4 years also display a positive association, although not statistically significant (p = 0.071052), indicating a potential impact.

4.8 Discussion of ZINB Model Fit Result

Using a multilevel count regression model based on EMDHS 2019 data, this study tried to examine the spatial distributions and determining factors of under-five mortality in Ethiopia. Under-five mortality serves as a vital health indicator and plays a crucial role in assessing human development (Rajaratnam, 2010). Its disparities within developing countries account for roughly half of the global gap between developed and developing countries. In this study, 5753 mothers who gave birth at least once in their life time were included. Of these, 906 (15.75%) women had lost at least one child by death before celebrating fifth birthday.

The spatial analysis in different methods consistently verified hot and cold spot areas of under-five mortality among under-five children in Ethiopia. The global test for spatial autocorrelation revealed 305 clustering in under-five mortality of child throughout the research areas without identifying the spatial location (moran's index = 0.2451, p-value<0.001). This could indicate that under-five mortality is clustered across EAs in Ethiopia.

Statistically significant hotspots for under-five mortality were identified in Benshangul Gumuz, Somali, the eastern part of SNNP, except SNNP this idea also supported by (Ali, 2022) as well as in the western part of Benshangul Gumuz and northern Gambela. Conversely, statistically significant coldspots, indicating lower risk, were found in Addis Ababa, central Oromiya, south Amhara, northern Tigray, and south Afar.

Through spatial kriging interpolation analysis, regions with a high risk of under-five child death were predicted. These high-risk areas were represented by red predictions. North Western Somali, Benshangul Gumuz, West Gambela, and western Oromiya were identified as regions with higher risk compared to others. On the other hand, the western part of South Amhara, Afar, Central Amhara, South and Central Oromia, and Addis Ababa were identified as less risky areas within the country.

Before the analysis of data using the multilevel approach, the heterogeneity of the under-five mortality with regard to regions was checked.

Consistent with the findings of (Gagabo, 2024), (SM, 2020), the results of this study revealed that there is a significantly lower expected number (or incidence rate) of under-five child death per mother when contraceptive methods are used compared to when they are not used. There is a significantly higher expected number (or incidence rate) of under-five child death per mother for birth orders (2nd and 3rd, 4th and 5th, 6 and above) and. multiple births compared to first birth and single birth, respectively which is in agreement with the findings of (Gagabo, 2024), (Tessema ZT, 2022). Larger household sizes (5-7 and above 7) are associated with a significantly lower expected incidence rate of under-five child death per mother compared to smaller household sizes (1-4), which is in line with the findings of previous studies (Habtamu, 2023), (Argawu, 2022).

In this study, an increase in the age of mothers was associated with an increase in the number (or incidence) of under-five mortality. Conversely, Ethiopia by (Bisrat Misganew Geremew, 2020). These studies revealed a correlation between a one-year increase in maternal age and a reduction in under-five mortality.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

The purpose of the study was to identify the determinants, assess EAs/cluster variations and describe the spatial distribution of the number of under-five child mortality in Ethiopia. The descriptive results showed that 75.54% of the mothers have not experienced under-five death , 15.75% of mothers experienced one under-five child death in their life time and only 0.07% of them lost 7 of their under- five children.

The spatial analysis verified hot and cold spot areas of under-five mortality among under-five children in Ethiopia. The global test for spatial autocorrelation revealed clustering in under-five mortality of children throughout the research areas without identifying the spatial location (Moran's I=0.416267, p-value<0.001). This indicates that under-five mortality is clustered across EAs in Ethiopia.

The identified statistically significant under-five mortality hotspot areas are Benshangul Gumuz, Somali, and the eastern part of SNNP, as well as the western part of Benishangul Gumuz and northern Gambela. . Conversely, Addis Ababa, central Oromia, south Amhara, northern Tigray, and south afar are identified as statistically significant cold spot regions (non-risky areas).

In the multilevel count regression analysis, individual mothers were considered as nested within the various enumeration areas in Ethiopia. There is an excess number of zeros, spatial auto covariate and unobserved heterogeneity (moderate intra class correlation) in the dataset. The test result suggested that the number of under-five death varies among enumeration areas and hence the multilevel count regression model fit better than the single level count regression model. The multilevel ZINB model fit results (count part) indicated that distance to get water, household size, sex house hold head, age of house hold head, age of mothers at the first birth, contraceptive use, birth order and type of birth have a statistically significant association with the number of under-five child mortality. The zero inflated part of the multilevel ZINB model fit result also revealed that excessive under-five deaths per mother are statistically significantly affected by season of child birth, region, toilet facility, house hold size (number of child in household), sex of household head, birth order of child, and birth interval.

From the zero-inflated results suggest that there may be regional variations in the factors influencing zero counts. Based on these findings, it would be valuable to investigate the specific factors or characteristics of these regions that may contribute to the lower odds of zero counts. Additionally, further research or data collection may be necessary to understand the underlying reasons for these regional differences and their implications for the population or outcomes being studied (Geremew , 2020).

5.2. Recommendations

Based on the findings of the study, we recommend the following actions:

Policy makers should develop targeted policies that consider the specific challenges and needs of different regions to address geographical variations in the number of under-five mortality. It is crucial to implement programs focused on reducing under-five mortality rates in areas with higher rates. Benshangul Gumuz, Somali, the eastern part of SNNP, the western part of Benshangul Gumuz, and northern Gambela need to be given special attention. Stakeholders, including government agencies, non-governmental organizations, and healthcare providers, should collaborate and coordinate their efforts to ensure a comprehensive approach to reducing under-five mortality. These measures can be extrapolated and applied to other regions to reduce disparities and improve the well-being of children and families. Such measures may include hospital delivery, family planning program and contraceptive use among others.

By supporting and empowering female household heads, we can work towards reducing under-five child mortality rates and improving the well-being of children in Ethiopia.

By increasing contraceptive use, individuals and families can make informed decisions about family size and spacing, leading to improved maternal and child health outcomes.

Our findings highlight the variations in the relationship between under-five child mortality and region, emphasizing the importance of considering regional factors in understanding and addressing child mortality rate.

References

(2023). Retrieved from https://en.wikipedia.org/wiki/Spatial_analysis.

A Bizzego, G Gabrieli, MH Bornstein, K Deater-Deckard, JE Lansford, RH Bradley, M Costa. (2021). Predictors of contemporary under-5 child mortality in low-and middle-income countries: A machine learning approach. *International journal of environmental research and public health*.

abbas Moghimbeigi & kazem Mohammad. (2008). Multilevel zero-inflated negative binomial regression modeling for over-dispersed count data with extra zeros. *Journal of Applied Statistics*.

Agidew, Belay. (2023). Spatial multilevel analysis of age at death of under-5 children and associated determinants: EDHS 2000–2016. *bmjopen.bmj.com*.

Agresti, A. (2012). *categorical data analysisSecond Edition* (Vol. Vol. 792). John Wiley & Sons,.

Ahmed, Kamal. (2016). Statistical analysis of factors affecting child mortality in Pakistan. *jcpsp.pk*, 543-544.

al., G. e. (2020). Factors Affecting Under-Five Mortality in Ethiopia using a multilevel negative binomial model. 11.

Alexandre Brouste. (2020). Closed form Maximum Likelihood Estimator for Generalized Linear Models in the case of categorical explanatory variables: Application to insurance loss modelling. *HAL open science*, 35, 689-724.

Ali, M. S. (2022). spatial variation and detrminants of under weight among children under 5 year of age in ethiopia. *nutration*, 102.

Amouzou A, e. a. (2018). Causes of child mortality: A systematic review. *Lancet, 391*, 1147-1157.

Andy H. Lee, K. W. (2006). Multi-level zero-inflated Poisson regression modelling of correlated count data with excess zeros.

Anselin, L. J. (1995). "Local indicators of spatial association—LISA. Geographical analysis,.

Argawu, A. S. (2022). Risk factors of under-five mortality in Ethiopia using count data regression models, 2021. *Annals of Medicine and Surgery*, 82.

Atalell KA, Alene KA. (2023). Spatiotemporal distributions of under-five mortality in Ethiopia between 2000 and 2019. (:. P. Jantchou, Ed.) *PLOS Glob Public Health*.

Awol M, E. D. (2023, February 16,). Spatial pattern and determinants of institutional delivery in Ethiopia: Spatial and multilevel analysis using 2019 Ethiopian demographic and health survey. (Abolfazl Mollalo, Ed.) *PLoS ONE*.

Ayele, T Zewotir. (2016). Childhood mortality spatial distribution in Ethiopia. *"journal of Applied Statistics"*, 1-17.

Bauldry, S. (2015). Bauldry, S. (2015). Bayesian Information Criterion. *scienceDirect*, 1-11.

Betts, M. G., Ganio, L. M., Huso, M. M., Som, N. A., Huettmann, F., Bowman, J., & Wintle, B. A. (2009). Comment on "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review". *Ecography*, 374-378.

Bevans, R. (2020). Akaike Information Criterion | When & How to Use It (Example). *scribbr*.

Bisrat Misganew Geremew, K. A. (2020). Factors Affecting Under-Five Mortality in Ethiopia: A Multilevel Negative Binomial Model, Pediatric Health, Medicine and Therapeutics. 525-534. Retrieved from https://doi.org/10.2147/PHMT.S290715

Bizzego, A. (2021). Predictors of contemporary under-5 child mortality in low-and middle-income countries: a machine learning approach. *environmental research and public sector, 18(3).*

Bosker, R., & Snijders, T. A. (2011). Multilevel analysis: An introduction to basic and advanced multilevel modeling. Multilevel Analysis,. 1-368.

Cameron A.C & P.K. Trivedi. (1986). Econometric models based on count data.Comparisons and applications of some estimators and tests. *applied econometrics 1*.

Collett, D. (2002). modelling binary data,CRC Press.

dash, s. k. (2023). brief introduction to multilevel modeling-analytics.

Dominique Lord & Byung-Jung Park. (n.d.). Negative Binomial Regression Models and Estimation Methods. 1-17.

EPHI and ICF. (2021). *Ethiopia Mini Demographic and Health Survey 2019: Final Report. Rockville, Maryland, USA: EPHI and ICF.* Ethiopian Public Health Institute (EPHI) [Ethiopia] and ICF, Addis Ababa.

F. Dormann. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review.

Fenta, Ayenew. (2020). The best statistical model to estimate predictors of under-five mortality in Ethiopia. *Journal of Big Data*.

Fenta, H. M. (2021). Spatial data analysis of malnutrition among children under-five years in Ethiopia. *BMC Medical Research Methodology*, 13.

GA Kaplan, JE Keil . (1993). Socioeconomic factors and cardiovascular disease: a review of the literature.

Gagabo, S. Y. (2024). Factors Influencing Mortality Under the Age of Five in Ethiopia. *World Journal of Public Health*, 17-27. Retrieved from https://doi.org/10.11648/j.wjph.20240901.13

Geremew . (2020). Factors Affecting Under-Five Mortality in Ethiopia: A Multilevel Negative Binomial Model. *An open access journal*, 11.

Getis, A. (2009). Spatial autocorrelation. . In *In Handbook of applied spatial analysis: Software tools, methods and applications* (pp. (pp. 255-278)). Berlin, Heidelberg: Springer Berlin Heidelberg.

Guddattu, V., Rao, K., & Rajkannan, T. (2015). Comparison between count regression and binary logistic regression models in the analysis of adverse drug reaction data. *In Probstat Forum, 08*, 140-147.

Habtamu, T. (2023). Determinants of Under-Five Child Mortality in Ethiopia: A Multilevel Count Analysis. *DSpace Institution's institutional repository*.

hilbe, j. (2011). Negative Binomial Regression, 2nd edn. Cambridge University Press,.

JA Tenedório, J Rocha. (2018). introductory chapter: spatial analysis modelling and planning. *3-15*.

JL., S. (1980). Illustrative analysis: infant and child mortality in Colombia. *Notas De Poblacion. 1980;8*.

JS Long and Freese. (2006). *Regression models for categorical dependent.*

Liu L, O. S. (2020). Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the Sustainable Development Goals. *338*, 3027-3035.

Liyew, A. M. (2021). Exploring spatiotemporal distribution of under-five mortality in Ethiopia: further analysis of Ethiopian Demographic and Health Surveys 2000, v. *BMJ Paediatrics Open*, 9.

McCullagh, P. . (UK (1989)). Generalized linear models. Routledge.

Meyer, J. (2008). *poisson or negative binomial? using count model diagnostic toselect model*. Retrieved from https://www.theanalysisfactor.com/poisson-or-negative-binomial-using-count-model-diagnostics-to-select-a-model/.

MM Fischer, J. W. (2011). *Spatial data analysis; models, methods and Techniques.*

NCSS.com. (n.d.). (NCSS.COM, Producer) Retrieved from https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Negative_Binomial_Regression.pdf.

Oloo. (2005). "Child mortality in developing countries: Challenges and policy . 1-17.

Overmars, K. D., De Koning, G. H. J., & Veldkamp, A. ((2003)). .Spatial autocorrelation in multi-scale land use models. . *Ecological modelling*,, 257-270.

Rajaratnam, J. K. (2010). Measuring under-five mortality : validation of new low-cost methods. *PLoS medicine*, 24.

S, A. (200). The determinants of infant mortality in Pakistan. *social science and medicens*, 51(2).

SB Javali, PV Pandit. (2010). Using zero inflated models to analyze dental caries with many zeroes:. *Indian Journal of Dental Research*.

science, E. c. (2024). *https://online.stat.psu.edu/stat504/lesson/9/9.1*.

SM, F. (2020). Risk factors of child mortality in Ethiopia: Application of multilevel two-part model. *PLoS ONE*, 15(8). Retrieved from https://doi.org/10.1371/journal.pone.0237640

Tessema ZT, T. T. (2022). Geographic variation and factors associated with under-five mortality in Ethiopia. A spatial and multilevel analysis of Ethiopian mini demographic and health survey 2019. (I. W. Alfredo Luis Fort, Ed.) *PLoS ONE*. Retrieved from https://doi.org/10.1371/journal.pone.0275586

Unicef. (2017). *Building the Future: Children and the Sustainable Development Goals in Rich Countries. Innocenti Report Card 14.* Innocenti, Florence: UNICEF Office of Research.

USAID. (2019). *Maternal, Neonatal and Child Health | Ethiopia. Retrieved from.* https://www.usaid.gov/ethiopia/global-health/maternal-and-child-health. Retrieved from https://www.usaid.gov/ethiopia/global-health/maternal-and-child-health.

Vuong, Q. H. (1989). "Likelihood ratio tests for model selection and non-nested hypotheses.".

WHO. (2019). *https://rho.emro.who.int/Metadata/under-five-mortality-rate*.

Wikipedia. (2024, April 25). Retrieved from https://en.wikipedia.org/wiki/Mortality_rate.

Wilson, P. (2015). The misuse of the Vuong test for non-nested models to test for zero-inflation. *ELSEVIER*.

Woldeamanuel, B. (2019). Socioeconomic, demographic, and environmental determinants of under-5 mortality in Ethiopia: evidence from Ethiopian demographic and health survey, 2016. *downloads.hindawi.com*.

Worku, M. G. (2021). Determinants of under-five mortality in the high mortality regions of Ethiopia: mixed effect logistic regression analysis. *Archives of Public Health*, 9.

zhoupeng ren, jinfeng and yilan liao. (2013). Using spatial multilevel regression analysis to assess soil type contextual.