Variations in Under-five Child Mortality among Regional States of Ethiopia: A Multi-level Modelling Approach

Ashenafi Senbeta Bedane[1*], Tilahun Ferede Asena[1,2], Aklilu Toma Shamenna[1] and Amanuel Disassa Abshoko[1]

[1]*School of Mathematical and Statistical Sciences, Hawassa University, P.O.Box 05, Hawassa, Ethiopia.*
[2]*Department of Statistics, Arba Minch University, P.O.Box 25, Arba Minch, Ethiopia.*

Authors' contributions

This work was carried out in collaboration between all authors. Author ASB designed the study, performed the statistical analysis, wrote the protocol, wrote the first draft of the manuscript and managed literature searches. Authors TFA, ATS and ADA managed the analyses of the study and literature searches. All authors read and approved the final manuscript.

Article Information

Original Research Article

ABSTRACT

Background: Improvements in child survival have been one of the major targets of development programs during the past three decades. The main aim of this study is to investigate under-five child mortality variations among regional states of Ethiopia. In this study, single level and multilevel binary logistic regression models is adopted for the analysis.
Methods: This study is conducted based on Demographic and Health Survey (DHS) 2011 data, collected for 10,156 children under-five years of age in Ethiopia.
Results: Based on the model adequacy tests the random intercept binary logistic regression model is found to be best fitting to the data. The variance of the random component model related to the intercept term is statistically significant, implying the presence of under-five child mortality variations among regional states of the country and it is accounted by the random intercept term.

_____

*Corresponding author: E-mail: ashenafi.senbeta@gmail.com;

Conclusion: The major significant factors affected under-five child mortality are: mother's education level, birth index, child size at birth, mother's age at birth, type of birth and breastfeeding status. It also revealed that there is a contribution of those major factors to under-five child mortality variations among regional states. However, those factors significantly affecting under-five child mortality is explicitly did not show significant effects on variations of under-five child mortality across regional states. The implication of the study is that all regional states should makes remedial measures on public health policy, design strategies to improve facilities and aptitudes of stakeholder living in their region toward those major factors affecting under-five child mortality and contributing to its variations among regional states to reduce under-five child mortality in the country.

1. INTRODUCTION

Improvements in child survival have been one of the major targets of development programs of Ethiopia during the past three decades. According to Ethiopian 2010 MDGs Reports trends and prospective for meeting MDG by 2015, sixteen out of every hundred children born in Ethiopia will not live beyond their fifth birth day [1]. Five will not even live the first month of life. Every year, approximately, 472,000 children under five years of age die in Ethiopia [1].

In 2000, the Ethiopia Demographic and Health Survey (EDHS) estimated the under-five mortality rate (U5MR) at 166 per 1000 live births. The Ethiopian Ministry of Health (MOH) estimates that the U5MR for 2002-03 was 140 [2]. In 2000, the rate of 166 placed Ethiopia at 21st in the world for under-five mortality [2]. Ethiopia neonatal mortality rate was relatively even higher and it was the fifth-highest in the world [2]. According to EDHS 2005, the rate of mortality in Ethiopia examined by comparing data from 2005 EDHS and 2000 EDHS, infant and under-five mortality rates, obtained for the five years preceding the two surveys, confirm a declining trend in mortality. Under-five mortality declined from 166 deaths per 1,000 live births in the 2000 survey to 123, while infant mortality declined from 97 deaths per 1,000 live births in the 2000 survey to 77 [2]. Even the child mortality has declined in Ethiopia still it is high child death at national level as compared to developed countries. Ethiopian MDG 4 would further reduction of child and maternal death.

The regional difference in child mortality may be due to differences in socio-economic composition [3], health-seeking behavior regarding child immunizations [4] and maternal and child health care utilization [4]. Indeed, the incorporation of community-level factors in the analysis of child mortality provides an opportunity to identify the health risks associated with particular social structures and community ecologies, which is a key policy tool for the development of public health interventions [5,6].

The number of under-five deaths in worldwide has declined from more than 12 million in 1990 to 7.6 million in 2010. Nearly 21,000 children under-five died every day in 2010 which was about 12,000 less than in 1990. Since 1990, the global under-five mortality rate has dropped 35 percent from 88 deaths per 1,000 live births in 1990 to 57 in 2010 [7]. The Northern Africa, Eastern Asia, Latin America and the Caribbean, South-eastern Asia, Western Asia and the developed regions have reduced their under-five mortality. The rate of under-five mortality was reduced in the year over 2000 to 2010, but remains insufficient to reach MDG 4, particularly in Sub-Saharan Africa, Oceania, Caucasus and Central Asia, and Southern Asia [7].

The highest rates of child mortality have been still in Sub-Saharan Africa where 1 in 8 children died before age five, which is more than the average for developed regions and Southern Asia [7]. Under-five mortality rates have fallen elsewhere and the disparity between these two regions and the rest of the world has grown [7]. Under-five deaths are increasingly concentrated in Sub-Saharan Africa and Southern Asia, while the share of the rest of the world dropped in 2010 [7].

A consistent series of estimates of under-five mortality rate provided historical trends, during the period of 1950–2000 for both developed and developing countries. On an average about 15% of newborn children in Africa are expected to die before reaching their fifth birthday [8]. The progress of infant and child mortality in Sub-Saharan Africa remained as a major health

problem, and the progress made during the past four decade has been unevenly distributed [9].

The neonatal, post-neonatal, infant and child mortality pattern are higher for mothers who are under 20 years of age. Infant and child mortality levels are lower for children whose mother's age is between 20 up to 29 [10]. Neonatal mortality of the children whose mothers aged is below 20 years at the time of the child's birth, is higher than the children whose mothers are in the age range 20-29 years at the time of giving birth [10]. Short birth intervals were significantly reduced infant probability of survival. The researchers used cross classification percentage distribution and logistic regression model [10].

An investigation on historical and modern third world countries have shown that children who are exclusively breast-fed survive longer and are healthier than artificially fed children in direct [11]. And also the breastfeeding practices have significantly lower risk among neonatal, post-neonatal and child mortality levels as compared to children never breastfed [10,12]. In Ethiopia a retrospective birth history data from a national survey used proportion hazard regression model [11] found a significantly higher risk of a conception in the months, following the death of an index child, even after controlling for breastfeeding status.

Maternal education has been identified as one of the most important socio-economic determinants of infant and child mortality. According study conducted in Sudanese in 2009 there are a significant association between mother's education and infant and child mortality. The researchers used statistical test of independence based on Chi-square [13]. There has been considerable decline in infant mortality as mother's educational attainment increases. However, there exists a marked differential among the regions. Infants born to women with no education are almost more likely to die before age one than infants born to women with primary and higher education and the researcher adopted multivariate logistic regression model in Ghana [14].

Improvements in child survival have been one of the major targets of development programs during the past three decades. A century later, out of the 187 countries, only nineteen countries- all in Africa- had an infant mortality rate of above ten percent. Ethiopia, through the progressive implementation of the Health Sector

Development Program in the last seven years, has made great strides to improve maternal and child survival. The reduction of infant and child mortality indirectly helps in reducing fertility by decreasing the desired number of children to be born due to increased probability of survival of a child. Under-five mortality is significantly influenced by breastfeeding status, ownership of toilet facilities, the level of education of the mother, residential area and place of delivery of the child; rural mothers and children are particularly at disadvantage with regards to basic health and socio-economic services based on logistic regression analysis and Cox regression [15]. Birth interval with previous child and mother standard of living index are the vital factor associated with child mortality. The cross-tabulation analysis [16] shows that birth interval with previous child and mother standard of living index is the vital factor associated with child mortality.

Regional disparities in under-five child mortalities are associated with factors at the community level that distinguish these regions from each other. The availability of services and social amenities in communities, or the lack infrastructure, may positively or negatively influence the health of the residents of communities. Some of these factors include differences in community-level development, population density, prevalence of poverty, and availability of maternal and child health care services. These are often interrelated aspects of the regional environment that are important for child health and well-being, and may also be relevant in exacerbating or mitigating inequities in resources and population health outcomes across regions [17,18].

The most recent studies related to child mortality in the regions within geographically diverse ecology and socioeconomic circumstances may have different disease exposures and child health outcomes. As [4] tried to assess variations in the risks of death in children under age 5 across regions of Nigeria and determined characteristics at the individual and community levels that explained possible variations among regions. The researcher applied multilevel Cox proportional hazards analysis using a nationally representative sample of 6,029 children from 2,735 mothers aged 15-49 years and nested within 365 communities from the 2003 Nigeria Demographic and Health Survey. Hazard ratios (HR) with 95% confidence intervals (CI) were used to express measures of association among

the characteristics. Variance partition coefficients and Wald statistic were used to express measures of variation. From the results, the researcher suggested the need to differentially focus on community-level interventions aimed at increasing maternal and child health care utilization and improving the socioeconomic position of mothers, especially in disadvantaged regions such as the South (Niger Delta) region [4].

Under-five children mortality in Ethiopia is one of the highest in the world and it is one of the challenging problems that the country needs to address. Even in an average year, the education, health and economic situation for millions of Ethiopian under-five children can only be described as a crisis.

1.1 Objectives of the Study

1) To identify the major factors that affect under-five child mortality and its variations among regional states of Ethiopia.
2) To assess within and between regional variations of under-five child mortality in Ethiopia.

2. DATA AND METHODOLOGY

Ethiopia is officially known as the Federal Democratic Republic of Ethiopia, is a landlocked country located in the Horn of Africa. It is the second-most populous nation in Africa, with over 82 million populations [2] and the tenth largest by area, occupying 1,100,000 km^2. Ethiopia is bordered by Eritrea to the North, Djibouti and Somalia to the East Sudan and South Sudan to the West, and Kenya to the South. Ethiopia has eleven geographic or administrative regions: nine regional states (Tigray, Affar, Amhara, Oromia, Somali, Benishangul-Gumuz, SNNPR, Gambela and Harari) and two city administrations (Addis Ababa and Dire Dawa that are considered as region) with capital city of Addis Ababa. Administratively, each of the 11 geographic regions in Ethiopia is divided into zones and each zone is divided into lower administrative units called *woredas*. Each *woreda* is then further subdivided into the lowest administrative unit, called a *kebele*.

2.1 Ethical Approval

This investigation was conducted according to the principles expressed in the Declaration of Ethiopian Statistical Association under the survey designed for Ethiopian Demographic and health survey (EDHS), Ethiopia. It was approved by the research ethics committee and all participants parents who agreed to participate in this study signed a consent form.

The Data: The sample for the 2011 EDHS designed to provide population and health indicators at the national and regional levels. The sample design allowed for specific indicators, such as contraceptive use, to be calculated for each of Ethiopia's eleven geographic / administrative regions: nine regional states and two city administrations. The sampling frame used for the 2011 EDHS was the Population and Housing Census conducted by the Central Statistical Authority (CSA) in 2007. During the 2007 PHC, each of the *kebeles* was subdivided into convenient areas called census enumeration areas (EAs). The 2011 EDHS sample was selected using a stratified, two-stage cluster design and EAs were the sampling units for the first stage. The 2011 EDHS sample included 624 EAs, 187 in urban areas and 437 in rural areas [2].

The data used for this study is 2011 Ethiopia Demographic and Health Survey (2011 EDHS). The survey was conducted under the guidance of the Ministry of Health by the Central Statistical Authority from 27, December 2010 through June 2011 with a nationally representative sample of nearly 18,500 households. But in this study, the data from Somali region was excluded from this study, because in the Somali region, in 18 of the 65 selected EAs listed households were not interviewed for various reasons, such as drought and security problems, and 10 of the 65 selected EAs, were not listed due to security reasons. Therefore, the data for Somali may not be totally representative of the region as a whole [2].

Variables of the Study: The variables considered in this study taken based on earlier studies at the global and national level. As discussed in the literature review socio-economic, demographic and environmental characteristics are to be the essential and proximate determinants of child mortality at worldwide and national level as well. In this study, the potential determinant factors expected to be correlated with under-five child mortality are included as variables. Those variables considered in this study are classified as: dependent and explanatory or indicator variables stated below.

Dependent Variable: The dependent variable of interest for this study is child event before reaching five years of age, measured as the duration from birth to the age at death. Since in the DHS age at death (reported in days and months) is subject to heaping at certain ages, a discrete formulation of time is preferred to a continuous one. It is dichotomous coded as 1 if child died in the five years before the survey and 0 if alive.

Explanatory Variables: In the present study the following socio-economic, demographic and environmental factors which are expected to have impacts on under-five child mortality in Ethiopia are classified as individual level variables and regional level variables.

2.2 Multilevel Logistic Regression Model

Before going to multilevel modeling, ones needed to go beyond the classical setup of a data Y and a matrix of predictors X. The multilevel data structures with an observational study of the impacts of each indicator of under-five child mortality. The treatment is at the groups (region) level, but the outcome is measured on individual families.

The fact that the regional states in Ethiopia had a variety of environmental factors, health service provider, level of education of the people living in the community, level of educated family, access to safe drinking water, sanitation and different infrastructures to encourage the reduction of under-five child mortality at their region and national level. Indeed, not only regional-level differentials but also there are the individual-level factors attributed for under-five child mortality in addition to demographic factors of children as well. This differential among individual, region, national and also through continent level indicated the facts that, the rate of child mortality in developed and developing country has different structure. But, so many studies in single level (eliminate those variation across regional states) regarding under-five child mortality in the world wide and at national level that invites errors. In fact, there is clear heterogeneity among the individual and regional-level characteristics that leads to variations while clustered those factors at single level.

In the present study, multilevel binary logistic regression model was adopted to model under-five child mortality variations among regional states of Ethiopia. This study, started to built multilevel modelling of the variations for the impacts of individual and community (regional)-level on under-five child mortality starting from empty, random intercept and random coefficient binary logistic regression model as discussed as follows. First, ones better to check where there is heterogeneity proportion of under-five child mortality between regions in Ethiopia before going to multilevel analysis.

2.2.1 Heterogeneity proportion

The basic data structure of the two-level regression is a collection of N groups ('units at two levels' or 'regions'), with in group j, $(j = 1, 2, \ldots, N$ random sample of n_j level-one units ('individual' or 'number of under-five children living in the region j').

Consider the outcome variable in equation (2.1), Y_{ij} (i = 1, 2, \ldots, n_j; j = 1, 2, \ldots, N) and denoted by for level-one unit i nested in level-two group j.

And the total sample size is $M = \sum_{j=1}^{N} n_j$. If one does not take explanatory variables into account, the probability of success is assumed constant in each group [19]. Let the probability of having under-five child death in region j be denoted by π_j. The dichotomous outcome variable for the child i in region j, Y_{ij} can be expressed as the sum of the probability in region j, π_j (the average proportion of i levels in region j, $E(Y_{ij}) = \pi_j$) plus some individual dependent residual, that is

$$ y_{ij} = \pi_j + \varepsilon_{ij} \tag{2.1} $$

The residual term is assumed to have mean zero and variance,

$$ \text{var}(\varepsilon_{ij}) = \pi_j (1 - \pi_j) $$

Since the outcome variable is coded 0 and 1, the group (region) sample average is the proportion of successes in group j given by:

$$ \hat{\pi}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} Y_{ij} $$

Where: $\hat{\pi}_j$ - is an estimate for the group-dependent probability $\hat{\pi}_j$. Similarly, the overall sample average is the overall proportion of successes, π and is given by:

$$\pi = \frac{1}{M}\sum_{j=1}^{N}\sum_{i=1}^{n_j} Y_{ij}$$

2.2.2 Test of heterogeneity proportion

For the proper application of multilevel analysis the first logical step is to test heterogeneity of proportions between groups. Here we present two commonly used test statistics that are used to check for heterogeneity [19]. To test whether there are indeed systematic differences between the groups, the well known Chi-Square test for contingency table can be used. In this case the Chi-Square test statistic is:

$$\chi^2 = \sum_{j=1}^{N} n_j \frac{(\hat{\pi}_j - \hat{\pi})^2}{\hat{\pi}(1-\hat{\pi})} \sim \chi^2(N-1) \quad (2.2)$$

It can be tested a chi-square distribution with $N-1$ degrees of freedom. This chi-squared distribution is an approximation valid if the expected number of success $(n_j\pi_j)$ (and of failures $(n_j(1-\pi_j))$ in each group all are at least one while 80 percent of them are at least 5 [20]. Estimations of Between and Within Group Variance: The true variance between the group dependent probabilities [19], i.e. the population values of π_j, is given by:

$$\hat{\tau}^2 = S_{between}^2 - \frac{S_{within}^2}{\tilde{n}} \quad (2.3)$$

Where: $\tilde{n}$ is defined as: $\tilde{n} = \frac{1}{N-1}\left\{M - \frac{\sum_{j=1}^{N} n_j^2}{M}\right\}$,

$$S_{between}^2 = \frac{\hat{\pi}(1-\hat{\pi})}{\tilde{n}(N-1)}X^2 \quad \text{and}$$

$$S_{within}^2 = \frac{1}{M-N}\sum_{j=1}^{N} n_j(1-\pi_j),$$

Where: χ^2 is given in equation (2.2)

The Empty Model: The empty two-level model for a dichotomous outcome variable refers to a population of groups (level-two units) and specifies the probability distribution for group-dependent probabilities π_j (probability of having i^{th} child in j^{th} group (region) died before five year of age) then, consider equation (2.1) without taking further explanatory variables into account. We focus on the model that specifies the transformed probabilities $f(\pi_j)$ to have a normal distribution. This is expressed, for a general link function $f(\pi)$, by the formula

$$f(\pi_j) = \beta_0 + U_{oj}$$

Where $f(\pi_j)$ - is the population average of the transformed probabilities β_o and U_{oj} is the random deviation from this average for group *j*. If $f(\pi)$ is the logit function, then $f(\pi_j)$ is just the log-odds for group *j*. Thus, for the logit link function, the log-odds have a normal distribution in the population of groups, which is expressed by:

$$Logit(\pi_j) = \beta_0 + U_{oj}$$

For the deviations U_{oj} it is assumed that they are independent random variables with a normal distribution with mean zero and variance δ_0^2. This model does not include a separate parameter for the individual level variance [19]. This is because the individual level residual variance of the y_{ij} (death or alive of under-five children follows Bernoulli distribution directly from the probability of having under-five child death (π_j) which is given by:

$$var(\varepsilon_{ij}) = \pi_j(1-\pi_j)$$

Denote by π_o the probability corresponding to the average value β_o, as defined by

$$f(\pi_o) = \beta_o$$

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

$$\pi_o = Logit\ (\beta_o) = \frac{\exp\ (\beta_o)}{1 + \exp\ (\beta_o)}$$

(2.4)

Because of the non-linear nature of the logit link function, there is no a simple relation between the variance of probabilities and the variance of the deviations U_{oj} [19]. According to [19] there is an approximate formula, however, valid when the variances are small. The approximate relation (valid for small δ_o^2) between the population variance is:

$$var(\pi_j) \approx \frac{\delta_o^2}{\left(f'(\pi_o)\right)^2}$$

For the logit function, this yields:

$$var(\pi_j) = \left(\pi_o(1 - \pi_o)\right)^2 \delta_o^2$$

Note that an estimate of population variance $var(\pi_j)$ can be obtained by replacing sample estimates of π_o and δ_o^2.

2.3 Random Intercept Binary Logistic Regression Model

With grouped data, a regression that includes indicators for groups is called a varying-intercept model because it can be interpreted as a model with a different intercept within each group [21]. In this case the random intercept model is consider only random effect of each indicators of under-five child mortality meaning that the region differ with respect to the average value of under-five child death , but there is no different relation between indicators of under-five child mortality among groups (regional states).

A assume that X is predictor's data matrix denoted by: $X_h, (h = 1, 2, ..., k)$ these variables are denoted by with their values indicated by x_{hij} [19]. Some or all of those variables could be level one variables, the success probability is not necessarily the same for all individual in a given group (region). From the above probability of having under-five child death depend on indicators was denoted by π_j. The outcome variable is split into an expected value and residual as in equation (2.1).

Then, random intercept model expresses the log-odds, i.e. the logit of π_{ij}, is the sum of a linear function of all indicators of under-five child mortality is given as:

$$Logit(\pi_{ij}) = \beta_{oj} + \beta_1\ x_{1ij} + \beta_2\ x_{2ij} + ... + \beta_k\ x_{kij} = \beta_{oj} + \sum_{h=1}^{k} \beta_h\ x_{hij}$$

(2.5)

Where, logit (π_{ij}) does not include a level-one residual because it is an equation for the probability of having under-five child death (π_{ij}) rather than for the outcome y_{ij}.

β_{oj} is assumed to vary randomly and is given by the sum of an average intercept β_o and group (region) dependent deviations U_{oj} is given:

By replacing $\beta_{oj} = \beta_o + U_{oj}$ in equation (2.5)

We have:

$$Logit\ (\pi_{ij}) = \beta_o + \sum_{h=1}^{k} \beta_h\ x_{hij} + U_{oj}$$

Or $$\pi_{ij} = \frac{\exp\left(\beta_o + \sum_{h=1}^{k} \beta_h\, x_{hij} + U_{oj}\right)}{1 + \exp\left(\beta_o + \sum_{h=1}^{k} \beta_h\, x_{hij} + U_{oj}\right)} \quad (2.6)$$

Where, β_h - is a unit difference between the X_h values of two individuals in the same group is associated with a difference of β_h in their log-odds, or equivalently, a ratio of $\exp(\beta_h)$ in their odds.

U_{oj} - is random part of the model and It is assumed that they are mutually independent and normally distributed with mean zero and variance δ_o^2.

2.3.1 Random slope binary logistic regression model

The multilevel modeling strategy accommodates the hierarchical nature of the DHS data and corrects the estimated standard errors to allow for clustering of observations within units [22]. A significant random effect may represent factors influencing the outcome variable that cannot be quantified in a large-scale social survey. A random effects model thus provides a mechanism for estimating the degree of correlation in the outcome that exists at the region level, while also controlling a range of all indicators may potentially influence the outcome.

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

$$\beta_{oj} = \beta_o + U_{oj}$$

$$\beta_{1j} = \beta_1 + U_{1j}$$

Thus, by substituting in equation (2.5) then, logit (π_{ij}) is given as:

$$Logit(\pi_{ij}) = (\beta_o + U_{oj}) + (\beta_1 + U_{1j})X_{1ij} = \beta_o + \beta_1 X_{1ij} + U_{oj} + U_{1j}X_{1ij} \quad (2.7)$$

Now, we have two random effects at group level, the random intercept U_{oj} and the random slope U_{1j}. It assumed that both random effects have mean zero. And the variances are denoted by δ_o^2, δ_1^2 and their covariance is δ_{o1}^2.

Where, β_o - is the average intercept of the response variable.

β_1 - is fixed regression coefficient given explanatory variable X_1.

U_0 - is the random coefficient in the model.

$U_o + U_1 X_{1ij}$ - is the random part of the model can be considered as interaction by group and predictors (X).

The two random effects that characterized group (region) U_{oj} and U_{1j} are correlated. Further, it is assumed that, for different groups, the pairs of random (U_{oj}, U_{ij}) effects are independent and identically distributed. Thus, the variances and covariance of the level-two random effects are (U_{oj}, U_{ij}) denoted by:

$$\text{var}\left(U_{oj}\right) = \delta_{oo} = \delta_o^2$$
$$\text{var}\left(U_{1j}\right) = \delta_{11} = \delta_1^2$$
$$\text{cov}\left(U_{oj}, U_{1j}\right) = \delta_{o1}^2$$

Now, we are going to extend the above single explanatory model by including more explanatory variable that has random effects on outcome variables. Suppose that there are k level-one explanatory variables $X_1, X_2, ..., X_k$, and consider the model where all predictor variables have varying slopes and random intercept.

That is:

$$\beta_{oj} = \beta_o + U_{oj}, \beta_{1j} = \beta_1 + U_{1j}, ..., \beta_{hj} = \beta_h + U_{hj}, \quad for\ h=1,2,...,k$$, then we have:
$$Logit\left(\pi_{ij}\right) = \left(\beta_o + U_{oj}\right) + \left(\beta_1 + U_{1j}\right)X_{1ij} + ... + \left(\beta_h + U_{hj}\right)X_{hij}$$
$$= \beta_o + \sum_{h=1}^{k}\beta_h X_{hij} + U_o + \sum_{h=1}^{k}U_{hj}X_{hij}$$

$$(2.8)$$

Where, $\beta_o + \sum_{h=1}^{k}\beta_h X_{hij}$ - is fixed part of the model and

$U_o + \sum_{h=1}^{k}U_{hj}X_{hij}$ - is the random part of the model

$U_{oj}, U_{1j}, ..., U_{hj}$ - are assumed to be independent between groups but may be correlated within groups. So the components of the vector $U_{oj}, U_{1j}, ..., U_{hj}$ are independently distributed as a multivariate normal distribution with zero mean vector and variances and co-variances Ω given by:

$$\Omega = \begin{bmatrix} \delta_o^2 & \delta_{1o} & . & . & . & \delta_{ko} \\ \delta_{o1} & \delta_1^2 & . & . & . & \delta_{k1} \\ \delta_{o2} & \delta_{12} & \delta_2^2 & & & \delta_{k2} \\ . & & & . & & . \\ . & & & & . & . \\ \delta_{ok} & \delta_{1k} & . & . & . & \delta_k^2 \end{bmatrix}$$

2.4 Multilevel Binary Logistic Regression Model Comparison

Deviance based on Chi-square: The deviance based on chi-square value for two models is obtained as two times the difference of log likelihood value of the two models. It is compared with the probability of deviance based on chi-square, is greater than critical value distributed to chi-squared at the difference between numbers of parameter in two models degree of freedom. If P-value is less than 5% level of significance, suggesting that multilevel empty model is significant.

The basic concept underlying this procedure is to compare the maximum likelihood under an assumed model with that of a baseline model.

Let $\hat{L}_c$ be the maximized likelihood under the current model. This statistic cannot be used on its own to assess the lack of fit of the current model unless compared with a corresponding statistic of an alternative baseline model for the same data. This latter model is taken to be a model that fits the data perfectly. Such a model will have the same number of unknown parameters as there are observations. The model is termed the full or saturated model and

the maximized likelihood under it is denoted by $\hat{L}_f$. The saturated model does not condense the information in the bulk of data into a simple summary, as it is not parsimonious. However, the maximum likelihood under this model is an intuitively appealing reference by which a corresponding value of a given model can be compared to assess the adequacy of the given model [19].

Let the statistic D, be defined as:

$$D=-2\log\left(\hat{L}_c \mid \hat{L}_f\right)=-2\left[\log\hat{L}_c -\log\hat{L}_f\right] \quad (2.9)$$

Large values of D are encountered when $\hat{L}_c$ is small relative to $\hat{L}_f$, indicating that the current model is a poor one. On the other hand, small values of D are obtained when $\hat{L}_c$ is similar to $\hat{L}_f$, indicating that the current model is a good one. The statistic D has chi-square distribution at degree of freedom equals to the difference between the number of parameter in full model and current model therefore, it measures the extent to which the current model deviates from the full model and is termed the deviance.

3. RESULTS

The under-five child mortality rate of Ethiopia is estimated at 69 per 1000 live births while, about 93.1% survived beyond their fifth birth day or five years of age. Likewise, the rate of under-five child mortality as shown in Table 1 the minimum death rates were observed at Addis Ababa which

is estimated to be 2.57% and the maximum death rates were observed at Benishangul-Gumuz region which is estimated to be 8.49%. Then after multiple covariates binary logistic regression and multilevel binary logistic regression analysis are used to identify the risk factors and to model the variations of under-five child mortality among regional states of Ethiopia, using 2011 EDHS data sets.

The likelihood ratio test of overall model shown in Table 2 indicates that there are significant relationships between under-five child mortality and its indicators. From Table 3, the Nagelkerke R square was found 25.6% indicating that, those indicators that had significant association with under-five child mortality included in binary logistic regression analysis are useful in predicting under-five child mortality and to indicate its variations among regional states.

Hosmer and Lemeshow test was found to be statistically insignificant as shown in Table 4. It indicates that, do not reject null hypothesis of the model fits the data very well. Thus, it indicates binary logistic regression model of under-five child mortality fits the Ethiopian Demographic and Health data very well.

3.1 Chi-square Test of Heterogeneity

The two-level structure is used with the region as the second-level unit and under-five children as level one unit. This is based on the idea that there may be differences in under-five child mortality between regions, which are not captured by the explanatory variables and hence may be regarded as unexplained variability within the set of all regions [19].

Table 1. Percentage of under-five child mortality death per region

Region	Number of children	U5CM	Death rate
Addis Ababa	388	10	2.57
Tigray	1148	71	6.18
Affar	1105	93	8.42
Amhara	1229	87	7.07
Oromiya	1690	112	6.63
Benishangul-Gumuz	895	76	8.49
SNNPR	1589	120	7.55
Gambela	815	63	7.73
Harari	633	41	6.47
Dire Dawa	664	31	4.67

Table 2. Likelihood ratio test of overall model

Model	-2*Log likelihood	df	χ^2	Overall df	P-value
Empty	5116.04	1	1086.4	28	0.000*
Full	4029.631	29			

Table 3. Model summary of binary logistic regression model

Step	-2*log likelihood	Cox & snell R square	Nagelkerke R square
1	4029.631	0.101	0.256

Table 4. Hosmer and lemeshow test of goodness fit

Step	Chi-square	df	P-value
1	12.134	8	0.145

(* Significant at 5%), (df - is degree freedom)

Table 5. Chi-square tests of heterogeneity

Chi-square tests			
Statistics	Value	df	P-value
Pearson Chi-square	27.058	9	0.001*

(*significant at 5% level)

Before attempting to multilevel analysis, one has to test the heterogeneity of under-five child mortality among regional states of Ethiopia. As it can be shown in Table 5, the Pearson Chi-square (χ^2_{calc}) = 27.058 which is greater than χ^2_{tab} = 16.918 at 9 degree of freedom with P-value = 0.001 which, is less than 0.05 level of significance, implying strong evidence of heterogeneity in under-five child mortality across regional states of Ethiopia. The estimated true standard deviation of under-five child mortality between and within-region were 0.014 and 0.9652 respectively. And also, the probability of under-five children died before their fifth birth day was found to be 0.069 (Appendix A: Table A1). Therefore, multilevel logistic regression model is attempted for modelling approach.

3.2 Multilevel Empty Model Analysis

The single level and multilevel empty (null) models, being compared, differ only with respect to the variance component and it is tested by the deviance of likelihood-ratio based on Chi-square as shown in Table 6. The probability of deviance based on Chi-square = 4.81 is greater than $\chi^2_{0.05(1)}$ = 3.841 or rejection probability

(P-value = 0.028), which is less than 0.05 level of significance (Table 6). Therefore, multilevel empty binary logistic regression analysis is found to be significant, suggesting overwhelming evidence of region effects on under-five child mortality in Ethiopia.

The overall mean of under-five child mortality is estimated at -2.64 and the between-region (level two) variance of under-five child mortality is estimated as δ^2_{uo}= 0.196 which is found significant at 5% level of significance, indicating the variations of under-five child mortality among regional states of Ethiopia was non-zero. The intra-region correlation of regional states of Ethiopia is estimated at 0.0115 which is found significant at 5% level of significance. It was indicating about 1.155% of the variance in under-five child mortality could be attributed to differences across regions.

3.3 Random Intercept Binary Logistic Regression Analysis

The variance of the random component related to intercept term is found to be significant. Indicating that, under-five child mortality variations among regional states of Ethiopia was non-zero. As shown in Table 7 mothers educational level, birth index, child size at birth, mother age at birth, type of birth and breastfeeding status have significant impacts and they have significant contribution to under-five child mortality variations among regional states of Ethiopia.

3.4 Random Coefficient Binary Logistic Regression Analysis

Based on predicted probability of under-five child mortality versus region effects of each variable those have significant impacts on under-five child mortality in the random intercept binary logistic regression analysis; birth index and child size at birth, are the indicators, supposed to be varying regionally. Thus, random slope binary logistic regression analysis of under-five child mortality is found insignificant at 5% significance level as shown in Table 8, suggesting no evidence that the major factors affects under-five child mortality

11

are explicitly did show an underline significant variations of under-five child mortality across the regional states of Ethiopia.

3.5 Multilevel Model Comparison

The deviance-based Chi-square test shown in Table 9 implied that, from binary multilevel logistic regression models, the random intercept binary logistic regression is significantly better fits the data. As compared to other multilevel models, the variations of under-five child mortality among regional states of Ethiopia for 2011 EDHS data sets was modeled by random intercept binary logistic regression. Again, the results based on AIC and BIC, which confirms the random intercept binary logistic regression model have less AIC and BIC, suggesting that random intercept binary logistic regression model as the best model for under-five child mortality variations among regional states of Ethiopia as compared to other multilevel models.

4. DISCUSSION OF RESULTS

The model comparison results depict random intercept binary logistic regression model is best fit the data very well. As a result of random intercept model mother educational level, birth index, child size at birth, mother's age at birth, type of birth and breastfeeding status are found to be covariates/incidences which have significant association to predict under-five child mortality variations across the region of Ethiopia. The overall under-five child mortality mean for this model is estimated to be 0.619 which is increased by about 3.26 as compared to empty multilevel model. And also, between-region (level two) variance of under-five child mortality is estimated as δ^2_{uo}= 0.218 which is increased by about 0.0218 as compared to the empty multilevel model. Thus, there is a significant

contribution of mother educational level, birth index, child size at birth, mother's age at birth, type of birth and breastfeeding status to under-five child mortality variations between regional states. This result is consistent with the earlier reports that said considerable regional differences in infant and child mortality rates in Ethiopia [23]. Again, the intra-region correlation is estimated at 0.014, implying that about 1.4% of the variations in under-five child mortality are attributed to differences between regions.

Based on the results of random intercept model, the odds of under-five child having mother with higher educational level being died before five years of age is reduced by 73.1% as compared to that having mother with no education. This result is consistent with the previous studies [12]. As a reason, it might be educated mothers are expected to have high awareness towards breastfeeding, higher gap between births (less birth index) and controlling the size of their children at birth as compared mothers with no education.

The probability of a child with higher birth index (less gap between births) being died before five years of age is higher than that of child born at least five year gap between its birth (birth index 1). This result is consistent with the previous study [11]. This might be, child born at least five year gap between its births (child with birth index 1) is well breastfed than that of child born less than five year gap in the same household.

Likewise, a child with large size at birth is more likely to die before its fifth birth day than under-five child with very small size at birth. Under-five child born from mother with age at birth from 25 up to 34 is less likely to die before five years age as compared to child born from mother with age at birth from 15 up to 24. This is consistent with the previous study [24].

Table 6. Results of multilevel empty binary logistic regression analysis

U5CM	B	Std. err.	Z	P>z	[95% CI. Interval]	
Fixed effect					Lower	Upper
β_0 = Intercept	-2.641	0.071	-33.8	0.000*	-2.79	-2.488
Random part						
Sigma(δ^2_{uo}) =Var	0.1962	0.086	2.27	0.023*	0.083	0.465
Intra-region correlation coefficient						
ICC (Rho (ρ_u))	0.0116	0.01	4.81	0.028*	0.002	0.062
Likelihood-ratio test of rho=0: chibar2(01) = 4.81						
Prob >= chibar2 = 0.014*						

(* Significant at 5%) and (ICC - intra-region correlation coefficient)

Table 7. Results of random intercept binary logistic regression analysis

Factors	Categories	OR [95% CI]	P-value
Fixed part			
Mother education level	Primary	0.938 [0.754, 1.166]	0.565
	Secondary	0.578 [0.307, 1.09]	0.091
	Higher	0.269 [0.088, 0.826]	0.022*
	No education (ref)	1	
Wealth index	Second lowest	1 [0.776, 1.29]	0.992
	Middle	0.943 [0.723, 1.229]	0.665
	Fourth	0.8 [0.604, 1.066]	0.129
	Higher	0.96 [0.635, 1.446]	0.84
	Lowest (ref)	1	
Birth index	Child 2	2.046 [1.696, 2.468]	0.000*
	Child 3	3.396 [2.466, 4.677]	0.000*
	Child 4	14.435 [6.122, 34.037]	0.000*
	Child with index 1 (ref)	1	
Child size at birth	Very larger	1.636 [1.24, 2.154]	0.000*
	Larger than average	1.759 [1.313, 2.357]	0.000*
	Average	0.985 [0.77, 1.26]	0.906
	Smaller than average	0.827 [0.565, 1.212]	0.331
	Very Smaller (ref)	1	
Month age at birth	25-34	0.643 [0.518, 0.798]	0.000*
	35-44	0.892 [0.696, 1.144]	0.369
	44+	1.271 [0.765, 2.112]	0.353
	15 up to 24 (ref)	1	
Type of Birth	Multiple	4.319 [3.12, 5.979]	0.000*
	Single (ref)	1	
Breastfeeding status	Breastfed	0.034 [0.026, 0.044]	0.000*
	Never breastfed	1	
Place of residence	Urban	1.022 [0.672, 1.551]	0.92
	Rural (ref)	1	
Intercept	$\beta_0 = 0.619$		0.003*
Random part	Sigma ($\delta^2_{ou} = 0.218$)		0.013*
			0.008*
	ICC (Rho (ρ_u)= 0.014)		

(Significant at 5%), (ref - is reference category), (ICC - intra-region correlation coefficient)*

Table 8. Results of random coefficient binary logistic regression analysis

Model diagnosis statistics	
-2*log likelihood	4053.748
Deviance based on Chi – Square	1.074
Degree of freedom (df)	27
P-value	0.956**
AIC	4107.747
BIC	4302.844

Table 9. Multilevel model comparison statistics

Model comparison statistics	Empty model	Random intercept	Random coefficient model
-2*log likelihood	5111.23	4054.82	4053.75
Deviance based on Chi – square	4.8138	1056.41	1.074
Degree of freedom (df)	2	22	27
P-value	0.0282*	0.000*	0.956
AIC	5115.23	4098.82	4107.75
BIC	5129.68	4257.79	4302.84

(Significant at 5% level)*

Similarly, the probability of child with multiple births is more likely to die than that of single birth. This result is consistent with the previous study [10]. And also, the probability of child having enough breastfed or high breastfed up to five years of age is less likely to die than under-five child never (less) breastfed up to five years of age. This result is consistent with the previous study [10].

5. CONCLUSIONS AND IMPLICATIONS OF THE STUDY

In Ethiopia under-five child mortality is significantly associated with geographical region. The probabilities of under-five children living in all regional states are more likely to die before five years of age than that of children living in Addis Ababa except children living Dire Dawa. Under-five child mortality variations among regional states were accounted by the random intercept terms of the model. Mother educational level, birth index, child size at birth, mother age at birth, type of birth and breastfeeding status are found to be significant incidence factors of under-five child mortality in Ethiopia. And also, these incidence factors have significant contribution in under-five child mortality variations among regional states. Based on random intercept model results, there is less probability of death before five years of age for children having educated mother, higher birth gap (less birth index), very small size at birth, mother with age at birth from 25 up to 34, single birth and enough breastfed till their fifth birth day.

The study suggests the following points to reduce under-five child mortality in Ethiopia. Government and non-governmental organization have to support mother's to educate themselves. Preferable if households have less birth index or less birth within five years. Improvement in maternal health care service will be appropriate to control larger size of child at birth. It is recommendable if mothers preferable to give birth at ages between 25 up to 34. Multiple born children need professional cares and special attention of their parents. Mothers have to develop the culture of breastfeeding of children. Further studies should be conducted to identify others factors that affect and contribute to under-five child mortality variations among regions. Multilevel models are appropriate method that investigates the effects of demographic, socio-economic and environmental factors on under-five child mortality and to take into account its variations among regional states.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. MoFED. Trends and prospects for meeting MDGs by 2015, MOH; Ethiopian MDGs Report, Addis Ababa, Ethiopia; 2010.
2. CSA. Ethiopia Demographic and Health Survey data. Central statistical authority (Ethiopia) and ORC Macro; Addis Ababa, Ethiopia and Calverton, Maryland, USA; 2011-2012.
3. Kandala NB, Ji C, Stallard N, Stranges S, Cappuccio FP. Spatial analysis of risk factors for childhood morbidity in Nigeria. Am J Trop Med Hyg. 2007;77:770-779.
4. Antai D, Ghilagaber G, Wedrén S, Macassa G, Moradi T. Inequities in under-five mortality in Nigeria: Differentials by religious affiliation of the mother. J Relig Health. 2011;48:290-304.
5. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: A critical review. J Epidemiol Community Health. 2001;55(2): 111-122.
6. Stephenson R, Baschieri A, Clements S, Hennink M, Madise N. Contextual influences on the use of health facilities for childbirth in Africa. American Journal of Public Health. 2006;96(1):84-93.
7. United Nation. The Millennium Development Report. Goals 4: Reduction of Child mortality, Addendum; 2011.
8. Omar BA, Lopez AD, Inoue M. The decline in child mortality: A reappraisal. Bulletin of the World Health Organization. 2000;78: 1175–1191.
9. Garenne M, Gakusi E. Health transitions in Sub-Saharan Africa: Overview of mortality trends in children under 5 years old (1950-2000). Bull World Health Organ. 2006; 84(6):470-478.
10. Mondal NI, Hossain K, Ali K. Factors influencing infant and child mortality: A case study of Rajshahi District, Bangladesh. J Hum Ecol. 2009;26(1): 31-39.
11. Lindstrom DP, Gebre-Egziabher K. The impact of infant and child death on subsequent fertility in Ethiopia. Institute of Public Health Florida A & M University Tallahassee; FL 32307 850-412-7422; 2001.

12. Tilahun Ferede Asena, Derbachew Asfaw Teni. Bayesian Semi-Parametric Regression Analysis of Childhood Malnutrition in Gamo Gofa Zone: The Social and Economic Impact of Child Under nutrition, American Journal of Theoretical and Applied Statistics. 2015; 4(4):269-276.

13. Mahfouz MS, Surur AA, Ajak DA, Eldawi EA. Level and determinants of infant and child mortality in Malakal town, Southern Sudan. Sudanese Journal of Public Health. 2009;4:2.

14. Goro IM. The Stalling child mortality in Ghana: the case of the three northern regions. University of Cape Town, Cape Town, South Africa; 2007.

15. Zeleke Worku. Factors That affect under-five mortality among South African children. Proceedings of the World Congress on Engineering and Computer Science, San Francisco, USA; 2009.

16. Kumar PP, Gemechis File. Infant and child mortality in Ethiopia: A statistical analysis approach, mathematics dept. Jimma Univerity, Indian Expatriate; 2010.

17. Siddiqi A, Irwin LG, Hertzman C. The region as a sphere of influence. Total environment assessment model for early child development: Evidence report World Health Organization's Commission on the Social Determinants of Health; 2007.

18. Tilahun Ferede Asena. Multilevel modelling of modern contraceptive use among Rural and Urban population of Ethiopia. American Journal of Mathematics and Statistics. 2013;3(1):1-16.

19. Snijders TAB, Bosker RJ. An introduction to basic and advanced multilevel modeling (1st edition). Department of Statistics, University of Poone; 1999.

20. Agresti A. Categorical data analysis. John Wiley and Sons, New York; 1990.

21. Gelman A, Hill J. Data analysis using regression and multilevel / hierarchical models. Columbia University, Journal of the American Statistical Association; 2006.

22. Goldstien H. Multilevel statistical models (4th edition). University of Bristol, UK. Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd; 2011.

23. Heins F, Mekonen Tesfaye, Valente P. Depth Studies from the 1994 Population and Housing Census in Ethiopia. Italian Multi-Bi Research Project ETH/92/P01; 2001.

24. Fikre Enquoselassie, Fottrell E, Byass P. The distribution and effects of child mortality risk factors in Ethiopia: A comparison of estimates from DSS and DHS, Ethiop. J. Health Dev. 2009;23:2.

APPENDIX A

Table A1. Estimation of between- and within-region variance of under-five child mortality

Region	n_j	$n_j * n_j$	$\hat{\pi}_j$	$(1-\hat{\pi}_j)$	$(n_j*\hat{\pi}_j)$ $*(1-\hat{\pi}_j)$	$(\hat{\pi}_j-\hat{\pi})$	$\dfrac{(\hat{\pi}_j-\hat{\pi})^2}{\hat{\pi}(1-\hat{\pi})}$	$n_j*\dfrac{(\hat{\pi}_j-\hat{\pi})^2}{\hat{\pi}(1-\hat{\pi})}$
					Estimation of between- and within-region variance			
Tigray	1148	1317904	0.06	0.94	1076.82	-0.007	0.0008	0.88
Affar	1105	1221025	0.08	0.92	1012.18	0.015	0.0035	3.87
Amhara	1229	1510441	0.07	0.93	1141.74	0.002	6E-05	0.08
Oromiya	1690	2856100	0.07	0.93	1578.46	-0.003	0.0001	0.24
Benzangul -Gumuz	895	801025	0.09	0.92	818.925	0.016	0.004	3.57
SNNPR	1589	2524921	0.08	0.92	1468.24	0.007	0.0008	1.21
Gambela	815	664225	0.08	0.92	752.245	0.008	0.001	0.81
Harari	633	400689	0.07	0.94	591.855	-0.004	0.0002	0.16
Addis Ababa	388	150544	0.03	0.97	377.912	-0.043	0.0288	11.2
Dire Dawa	664	440896	0.05	0.95	632.792	-0.022	0.0075	5
Total	10156	11887770			9451.17			27

(π - is dependent probability proportions for each region)

Where, $\pi=\dfrac{Total\ under-five\ yeras\ of\ age\ child\ mortality}{Total\ number\ of\ under-five\ years\ of\ age\ children}=\dfrac{704}{10156}=0.069$

$$\chi^2 = \sum_{j=1}^{N} n_j \frac{(\hat{\pi}_j - \hat{\pi})^2}{\hat{\pi}(1-\hat{\pi})} = 27$$

$\tilde{n}$ is defined as: $\tilde{n} = \dfrac{1}{N-1}\left\{M - \dfrac{\sum_{j=1}^{N} n_j^2}{M}\right\} = \dfrac{1}{10-1}\left\{10156 - \dfrac{11887770}{10156}\right\} = 998.387$

$$S^2_{between} = \frac{\hat{\pi}(1-\hat{\pi})}{\tilde{n}(N-1)}X^2 = \frac{0.69(1-0.69)}{988.387*(10-1)}*27 = 0.00193$$

$$S^2_{within} = \frac{1}{M-N}\sum_{j=1}^{N} n_j \pi_j (1-\pi_j) = \frac{9451.2}{10156-10} = 0.9315$$