International Journal of Multidisciplinary Research and Development
Online ISSN: 2349-4182, Print ISSN: 2349-5979; Impact Factor: RJIF 5.72
Received: 27-12-2019; Accepted: 29-01-2020
www.allsubjectjournal.com
Volume 7; Issue 2; 2020; Page No. 197-207

Applications of count data models for determinants of under-five mortality in Ethiopia

Tegegn Terefe[1] Tesfaye Denano[2], Sintayehu Sibera[3]
[1-3] Wolaita Sodo University, Department of Statistics, POBO 138, Sodo, Ethiopia

Abstract
Under-5 mortality is risk of a child dying before completing five years of age. The purpose of this study was to identify factors that affect under-five mortality based on 2016 EDHS data set using count regression model with specific objectives: to examine socio economic and demographic factors in terms of under-five death and to identify the appropriate count regression model to fit under-five death data. The survey collected information from a total of 15,683 women aged 15-49 years out of which 11951 women are considered in this study. The data were found to have excess zeros (62.5%) and the variance was higher than its mean. Thus several count models such as Poisson, NB, ZIP, ZINB, HP and HNB models were fitted to select the best model which fits the data. Among these models HNB was more appropriate to fit the data according to model selection criteria. Based on HNB model result the truncated negative binomial part, the explanatory variables like region, mother's education level, contraceptive use, birth order, breastfeeding and birth type are statistically significant factors while for logistic part, using breastfeeding method, place of residence, region, mother's occupation, mother's educational level, contraceptive use, wealth index, birth type and birth order are statistically significant factor on the number of under-five death per mother in Ethiopia. Hence, the concerned body should work closely with both the private sector and civil society to teach households to have sufficient knowledge and awareness on under five mortality and mechanisms of reduction and moreover to make children very well.

Keywords: count data models, under-five mortality, AIC, Ethiopia

1. Introduction

Over the last two decades, the world made substantial progress in reducing mortality among children (including children under age 5, children aged 5-9 and young adolescents aged 10-14). Still, in 2017 alone, an estimated 6.3 million children and young adolescents died, mostly from preventable causes. Children under age 5 accounted for 5.4 million of these deaths, with 2.5 million deaths occurring in the first month of life, 1.6 million at age 1-11 months, and 1.3 million at age 1-4 years [23].

Children continue to face widespread regional and income disparities in their chances of survival. Sub-Saharan Africa remains the region with the highest under-five mortality rate in the world. In 2017, the region had an average under-five mortality rate of 76 deaths per 1,000 live births. This translates to 1 in 13 children dying before his or her fifth birthday. This is 14 times higher than the average ratio of 1 in 185 in high-income countries and 20 times higher than the ratio of 1 in 263 in the region of Australia and New Zealand, which has the lowest regional under-five mortality rate. The world has made substantial progress in reducing child mortality in the past several decades, and millions of children have better survival chances than in 1990. That is 1 in 26 children died before reaching age 5 in 2017, compared to 1 in 11 in 1990. Moreover, progress in reducing child mortality has been accelerated in the 2000–2017 period, compared with the 1990s, with the annual rate of reduction in the global under-five mortality rate increasing from 1.9 percent in 1990–2000 to 4.0 percent in 2000–2017. Despite the global progress in reducing child mortality over the past few decades, an estimated 5.4 million children under age five died in 2017 and roughly half of those deaths occurred in sub-Saharan Africa. Globally, in 2017, half of all deaths under 5 years of age took place in sub-Saharan Africa, and another 30% in Southern Asia. In sub-Saharan Africa, 1 in 13 children died before their fifth birthday. In high-income countries, that number was 1 in 185 [27].

One of the objectives of the Sustainable Development Goals (SDG) is to diminish under-five mortality rate and improvement in maternal health which by implication increases the chance of child survival. According to UNCEF report, poverty is one of the most important factors affecting the under- five mortality rate in Africa, including Ethiopia [28]. Under-five mortality rates dropped by more than half from 1990 to 2016, decreasing from an estimated 93 deaths per 1000 live births to 41 deaths per 1000 live births, respectively. Much of that reduction has been achieved in recent years, with the rate of decline increased from 1.9 percent in 1990-2000 to 4.0 percent in 2000-2016. Even though there is reduced; child survival remains a serious concern globally [26].

Despite remarkable success, child survival remains a serious concern globally. Unacceptably, about 15,000 children still die every single day globally. The level of under-five mortality remains high in certain regions of the world, especially in Southern Asia and sub-Saharan Africa. Sub-Saharan African Region continues to be the region with the highest rate of under-five mortality. For instance, in 2016 the under five-mortality rate in sub-Saharan African Region was the highest in the world, 79 deaths per 1,000 live births, nearly 15 times the average in developed countries[26].

Under-five mortality in Ethiopia is one of the highest in the world and 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. Factors such as low level of mother's education, unsafe drinking water and sanitation, low family income, birth interval, short to breast feeding time, lack of place of birth delivery and periodic famine continue to put children at risk. According to 2011 EDHS, one in 17 Ethiopian children dies before the first birthday and one in 11 Ethiopian children dies before the fifth birthday. An Ethiopian child is 30 times more likely to die by his or her fifth birthday than a child in Western Europe. These levels of child mortality in Ethiopia are higher than the target of the minimum Millennium Development Goals (MDGs) which are 88 deaths per 1000 live births [12].

The 2016 EDHS results show that the neonatal, infant and under-five mortality rates for the five years before the survey are 29, 48, and 67 deaths per 1,000 live births, respectively. In other words, in Ethiopia 1 in every 35 children dies within the first month, 1 in every 21 children dies before celebrating the first birthday, and 1 of every 15 children dies before reaching the fifth birthday. By the end of MDG era, 2030 agenda for sustainable development was presented. Seventeen Sustainable Development Goals (SDGs) were agreed by global leaders on the basis of millennium development goals. The target is to drop the "neonatal mortality as low as 12 per 1000 live births and under-five Mortality to as low as 25 per thousand live births" [28].

In recent years, child mortality in Ethiopia is reported to show a decrement. For instance, neonatal mortality decreased by 41 percent over the 16-year period between 2000 and 2016 EDHS, from 49 deaths per 1,000 live births to 29 deaths per 1,000 live births. Infant mortality declined by 50 percent during the same period (declined from 97 deaths per 1,000 live births to 48 deaths per 1,000 live births). Under-five mortality also decreased by 60% during the same period (declined from 166 deaths per 1,000 live births to 67 deaths per 1,000 live births). The decline reported being continued in the recent EDHS with neonatal, infant and under-five mortality of 29, 48 and 67 deaths per 1,000 live births, respectively [9]. Even if the mortality rate is decreased from time to time, it remains to be among the highest in the world. As a result, about one in every 35 Ethiopian children dies before his/her first month, one in every 21 children dies before the first birthday, and one in every 15 children dies within the fifth birthday. Therefore, the focus of this study was investigating the magnitude of childhood mortality and its predicting factors in Ethiopia to fill the gaps that existed in the country.

2. Method of Data Analysis

In this study, the variable of interest is a count variable. When the response or dependent variable is a count (which can take on non-negative integer values (0, 1, 2…), it is appropriate to use non-linear models based on non- normal distribution to describe the relationship between the response variable and a set of predictor variables.

Count regression models were developed to model data with integer outcome variables. These models can be employed to examine the occurrence and frequency of occurrence. The most popular model for count data is the Poisson model, which is based on the property that the mean and variance of the dependent variable is assumed to be equal [15]. However, this is not always the case, as the variance sometimes exceeds the mean. Over dispersion can be modeled by using negative binomial (NB) regression model, but more models accounting for over dispersion exist. The negative binomial regression model assumes a gamma distribution for the Poisson mean with variation over the subjects.

2.1 General count data models

2.1.1 Poisson Regression Model

The Poisson distribution was developed to model discrete counts, and because it is similar to linear regression in many respects, it is relatively easy to interpret. This distribution becomes increasingly positively skewed as the mean of the dependent variable decreases [19], reflecting a common property of count data. The apparent simplicity of Poisson comes with two restrictive assumptions. First, the variance and the mean of the count variable are assumed to be equal. In reality, however, the variance is usually much greater than the mean which is called over-dispersion [8] and therefore, Poisson models, though widely used to handle count data may not be well suited to handle some types of count outcomes. Another restrictive assumption of Poisson models is that occurrences of the event are assumed to be independent of each other. This assumption is also frequently violated. The Poisson probability model is appropriate for events that occur randomly over time and/or space. The Poisson probability mass function, with rate parameter μ, is given by:

$$p(Y = y_i) = \frac{\exp(-\mu)\mu^{y_i}}{y_i!} \quad , \ \mu > 0 \ \text{and} \ i = 0,1,2,3,... \tag{1}$$

Where, y_i is the number of under-five deaths for i^{th} mothers in a given time with rate parameter μ, the mean and variance of the Poisson distribution is given as

$$E(Y) = Var(Y) = \mu \tag{2}$$

To proceed, we assume that for each individual mother, the probability of her children dying depends on the number of children exposed to the risk of mortality, hence children ever born. This then allows us to control the number of children exposed to the risk for a given woman, which we call an offset. Let $\ln(N_i)$ be an offset, where N_i is defined as the total number of children a mother had in her lifetime (the total number of live births from i^{th} mother).

Our interest focuses on how the mean number of events changes due to changes in one or more of the factors. Suppose that we want to consider k factors, $X_1, X_2, ... X_k$ simultaneously. The expected value of N_i can be written as:

$$E(Y_i) = N_i \exp(\beta_O + \sum_{j=1}^{k} \beta_j X_{ji}) \quad (3)$$

Taking natural logarithms, this is equivalent to:

$$\ln(\mu_i) = \ln(N_i) + \beta_O + \sum_{j=1}^{k} \beta_j X_{ji} \quad (4)$$

Where, β is k+1 dimensional parameter vector affecting under-five mortality levels and $\exp(\beta_O + \sum_{j=1}^{k} \beta_j X_{ji})$ is the risk of the i^{th} mother

The log-likelihood function is

$$l(\mu, y) = \sum_{i=1}^{n} \left(y_i \log(\mu_i) - \mu_i - \log(y_i!) \right) \quad (5)$$

Estimation of parameter is obtained by taking the partial derivations of the log-likelihood function and setting it equal to zero.

Over dispersed (Extra-) Poisson model

It is possible to account for over dispersion with respect to the Poisson model by introducing a dispersion parameter

$$\text{var}(y_i) = \alpha\mu \quad (6)$$

When $\alpha = 1$(near to 1) we have an ordinary Poisson model, when $\alpha > 1$we have the over dispersed Poisson model. Note that the introduction of the dispersion parameter, however, does not introduce a new probability distribution, but gives a correction term for testing the parameter estimates under Poisson model. The models are fit in the usual way, and the parameter estimates are not affected by the value of α, but the estimated covariance matrix is inflated by this factor. This method produces an appropriate inference if over dispersion is modest and it has become the conventional approach in the Poisson regression analysis. There are two ways of dealing about over dispersion: 1) Adjust estimated standard errors; when we are primarily concerned with testing hypotheses regarding parameter estimates. That is multiplying the standard errors of all the coefficient estimates by the square root of the estimated over dispersion. Without this adjustment, the confidence intervals would be too narrow, and inferences would be overconfident. 2) Use an alternative distribution as your random component (i.e., model the extra variability) — When our concern is prediction.

Test for over-dispersion

A natural basis for testing the adequacy of the Poisson model is the relationship between $\text{var}(y_i \mid x_i, \beta)$ and $\text{E}(\text{y}_\text{i} \mid \text{x}_\text{i}, \beta)$. To assess the adequacy of the negative binomial model over the Poisson regression model, we test the hypothesis that the mean and the variance are equal (equi-dispersion). Here the diagnostic tests are concerned with checking for this assumption. Deviance and Pearson Chi-Square divided by the degrees of freedom are used to detect over dispersion or under dispersion in the Poisson regression. Values greater than 1 indicate over dispersion, that is, the true variance is bigger than the mean, whereas values smaller than 1 indicate under dispersion, that is, the true variance is smaller than the mean. Evidence of under dispersion or over dispersion indicates inadequate fit of the Poisson model [20]. According to [3], as a rule of thumb the Deviance divided by the degree of freedom should be approximately equal to one for a satisfactory model. We can also test dispersion by specifying the extra Poisson variation. Here the extra Poisson parameter is estimated; say as α instead of 1, to give the variance a $\alpha\mu$ i.e $\text{var}(\text{y}_\text{i} \mid \mu_\text{i}) = \alpha\mu_\text{i}$. This α value is used to detect the over dispersion or under dispersion in the Poisson regression. α Value greater than 1 indicate over dispersion whereas α value smaller than 1 indicate under dispersion. Therefore, in testing over dispersion, the hypothesis is given by:

$$\text{H}_\text{o} : \alpha = 1 \quad \text{againt} \quad \text{H}_\text{a} : \alpha > 1$$

It is suggested that the over dispersion parameter α can be estimated by:

$\hat{\alpha} = \frac{\chi^2}{(n-k)}$, where χ^2 is the Pearson‘s chi-square statistic, n is the number of observations, and k+1 is the number of unknown regression parameters in the Poisson model. The Pearson‘s chi-square statistic is given by:

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \tag{7}$$

where $\hat{\mu}_i = N_i \exp(X_i \hat{\beta})$

2.1.2 Negative Binomial Regression Model

The negative binomial regression model is more flexible than the Poisson model and is frequently used to study count data with over-dispersion [14]. In fact, the negative binomial regression model is in many ways equivalent to the Poisson regression model because the negative binomial model could be viewed as a Poisson-gamma mixture model [14]. However, the difference is that the negative binomial regression model has a free dispersion parameter. In other words, the Poisson regression model can be considered as a negative binomial regression model with an ancillary or heterogeneity parameter value of zero. In the negative binomial regression model, a random term reflecting unexplained between-subject differences is included [10], that is, the negative binomial regression adds an over-dispersion parameter to estimate the possible deviation of the variance from the expected value under Poisson regression. Therefore, using the negative binomial regression to model count data with a Poisson distribution has the consequence of generating more conservative estimates of standard errors and may modify parameter estimates. Negative Binomial (NB) distribution is a way of modeling over-dispersed count data for $y_i \mid x_i$ which can arise as gamma mixture of poison distribution. One parameterization of its probability mass function is

$$P(Y_i = y_i) = \frac{\Gamma(y_i + \alpha)}{\Gamma(y_i + 1)\Gamma(\alpha)} \frac{\alpha^{\alpha} \mu_i^{*y_i}}{(\alpha + \mu_i^*)^{\alpha + y_i}}, \quad \mathrm{i} = 0,1,2,... \tag{8}$$

Where μ_i^* is the mean of Y_i and α is the inverse dispersion parameter, which is defined as:

$$\ln \mu_i^* = X_i \beta + \varepsilon_i \tag{9}$$

Where, $\mu_i^* = E(Y_i) = \exp(X_i \beta)$, and $\mathrm{var}(\mathrm{Y_i}) = \mu_i^* + \alpha \mu_i^{*2}$

Note that a Poisson random variable is a special case of a negative binomial random variable when α is allowed to become infinite. This is further evidence of the flexibility of the negative binomial distribution since there are infinitely many other choices for α that yield something other than a Poisson distribution.

2.2 Zero-Inflated Count Regression Models

The NB model might not be appropriate if the over-dispersion is caused by an excessive number of zeros in the outcome. In these cases, alternative models such as zero inflated models are recommended. The ZIP model, introduced by [18], is served as a dual-state method for modeling data characterized by a significant amount of zeros or more zeros than the one would expect in a traditional Poisson or negative binomial model, while the ZINB model, introduced by [14], is a more flexible model that can be used to handle over-dispersion caused by both unobserved heterogeneity and excess zeroes. Zero-inflated regression also considers two data generating processes and it assumes zero counts come from two different sources. Specifically, a zero count may come from the always-zero group (mothers who are never born) or the not always-zero group (mothers who may not be dead her child). Zero-inflated regression is also a two-part model. A Logit model determines if a zero count is from the always-zero group or the not-always-zero group and a baseline model, either Poisson or Negative binomial, governs both zero and positive counts from the not-always-zero group. A ZIP will reflect data accurately when over-dispersion is caused by excess of zeros. If over dispersion is attributed to factors beyond the inflation of zeros, a ZINB model is more appropriate [19].

2.2.1 Zero-Inflated Poisson (ZIP) Regression Model

ZIP model, well described by [18] is a simple mixture model for count data with excess zeros. The model is a combination of a Poisson distribution and a degenerate distribution at zero. Specifically, if Y_i is the number of under-five mortality per mothers is independent random variables having a zero-inflated Poisson distribution, the zeros are assumed to arise in two ways corresponding to distinct underlying states. The first state occurs with probability π_i and produces only zeros (mothers who are never born), while the other state occurs with probability $1 - \pi_i$ and leads to a standard Poisson count with mean μ and hence a chance of further zeros (mothers who may not be dead her child). In general, the zeros from the first state are called structural zeros and those from the Poisson distribution are called sampling zeros [16]. This two-state process gives a simple two-component mixture distribution with probability mass function

$$p(Y_i = y_i) = \begin{cases} \pi_i (1-\pi_i)\exp(-\mu_i) & \text{if } y_i = 0 \\ (1-\pi_i)\dfrac{\exp(-\mu_i)\mu_i^{y_i}}{y_i!} & \text{if } y_i = 1,2,..... \end{cases} \qquad 0 \le \pi_i \le 1 \tag{10}$$

The parameter μ_i and π_i depends on the covariates x_i and z_i, respectively. The mean and the variance of ZIP regression model, respectively, are:

$$E(y_i) = (1-\pi_i)\mu_i \quad \text{and} \quad \mathrm{Var}(y_i) = \mu_i(1-\pi_i)(1+\pi_i\mu_i)$$

Indicating that the marginal distribution of y_i exhibits over-dispersion, if $\pi_i > 0$. It is clear that this reduces to the standard Poisson model when $\pi_i = 0$. For positive values of π_i we have zero inflation, however, it is possible for $\pi_i < 0$ and to still obtain a valid probability distribution (this corresponds to a deficit of zeros (zero-deflation) [16].

2.2.2. Zero-Inflated Negative Binomial (ZINB) Regression Model

However, the ZIP model may often fail to fit such data either because of over-dispersion in relation to the Poisson distribution. We extend the ZIP regression model to ZINB regression model. The ZINB regression is used for count data that exhibit over-dispersion and excess zeros [19, 30]. The over-dispersed data are characterized by "excess zeros", "excess large outcomes" or both. ZINB model, therefore accounts for "excess zeros" and also for extra heterogeneity in a positive outcome. The ZINB distribution is a general model for counts which nests the ZIP, NB, and Poisson models [8].

The ZINB regression model assumes there are two distinct data generation processes. The result of a Bernoulli trial is used to determine which of the two processes are used. For observation *i*, with probability π_i the only possible response of the first process is zero counts, and with the probability of $1-\pi_i$ the response of the second process is governed by a negative binomial with mean μ *i*.

The zero counts are generated from the first and second processes, where a probability is estimated for whether zero counts are from the first or the second process. The overall probability of zero counts is the combined probability of zeros from the two processes. ZINB also arises in Bernoulli trials with non-equal success probabilities.

Suppose Y_i is the number of under-five mortality per mothers, then the probability mass function of ZINB is given by:

$$p(Y_i = y_i) = \begin{cases} \pi_i + \dfrac{(1-\pi_i)}{(1+\alpha\pi_i)^{-\frac{1}{\alpha}}}, & \text{if } y_i = 0 \\ (1-\pi_i)\dfrac{\Gamma(y_i + \frac{1}{\alpha})}{y_i!\Gamma(\frac{1}{\alpha})}(1+\alpha\mu_i)^{-\frac{1}{\alpha}}(1+\dfrac{1}{\alpha\mu_i})^{-y_i}, & \text{if } y_i > 0 \end{cases} \qquad 0 \le \pi_i \le 1 \tag{11}$$

Where, μ_i is mean of the underlying negative binomial distribution and α is the over dispersion parameter. The ZINB distribution reduces to the ZIP distribution as $\alpha \to 0$. The ZINB model is a special case of a two-class finite mixture model like the ZIP model with mean and variance, $E(y_i) = (1-\pi_i)\mu_i =$ and $\mathrm{Var}(y_i) = (1-\pi_i)(1+\pi_i\mu_i + \alpha\mu_i)$ respectively [30].

2.3 Hurdle Models

The development of the hurdle model is driven by the motivation to solve the problem of excessive number of zeros. This refers to observing more zeros than predicted by baseline count models such as Poisson or Negative Binomial. The hurdle regression handles the excess zeros by relaxing the assumption that zeros and positives come from a single data generating process [21]. The hurdle model is flexible and can handle both under and over dispersion problems. A hurdle model is introduced by [13] for the analysis of over dispersed or under-dispersed count data. The hurdle model approach, like the ZI model approach, is a 2-part count regression method that deals with the phenomenon of excess zeros in the data. However, hurdle models are distinct from ZI models. The first component of a hurdle model is typically a binomial distribution, determining if a count is zero or positive (as opposed to an excess zero') so that it pertains to prevalence (or incidence) in the overall population, as it targets all zero counts. The second part is a truncated at zero models governing the positive counts, i.e. $E(Y_I / Y_I > 0)$ [8].

2.3.1 The Poisson Hurdle Model

The Poisson Hurdle model is a two-component model comprising of a hurdle component models zero versus non-zero counts, and a truncated Poisson hurdle component is employed for the non-zero counts [8, 13].

$$p(Y_i = y_i) = \begin{cases} \pi_o & \text{if } y_i = 0 \\ (1-\pi_o)\dfrac{\exp(-\mu_i)\mu_i^{y_i}}{(1-\exp(-\mu))y_i!} & \text{if } y_i = 1,2,3,\ldots.. \end{cases} \quad 0 \le \pi_o \le 1 \tag{12}$$

$$where \quad \pi_o = p(y_i = 0 \quad \text{and} \quad \mu_i = E(x_i\beta)$$

2.3.2. The Negative Binomial Hurdle Model

Similarly, for the hurdle models, the Negative Binomial Hurdle can be used instead of Poisson distribution above in case of over-dispersion [13]. We consider a hurdle negative binomial (HNB) regression model in which the response variable has the Y_i = (i =1, 2, 3, n) has the distribution

$$p(Y_i = y_i) = \begin{cases} \pi_o & \text{if } y_i = 0 \\ (1-\pi_o)\dfrac{\Gamma(y_i + \frac{1}{\alpha})(1+\alpha\mu_i)^{-\frac{1}{\alpha}}(1+\frac{1}{\alpha\mu_i})^{-y_i}}{y_i!\Gamma(\frac{1}{\alpha})\left(1-(1+\alpha\mu_i)^{-\frac{1}{\alpha}}\right)} & \text{if } y_i > 0 \end{cases} \quad 0 \le \pi_i \le 1 \tag{13}$$

Where, $\alpha \ge 0$ is a dispersion parameter that is assumed not to depend on covariates? In addition, we suppose $0 < \pi_0 < 1$ and $\pi_0 = \pi_0(z_i)$ satisfy.

Parameter Estimation

The most commonly used methods of estimating the parameter of a count regression model is the method of maximum likelihood estimation (MLE). The method maximum likelihood parameter estimation is to determine the parameters that maximize the probability (likelihood) of the sample data. The log likelihood functions ofβ,

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

$$= \sum l_i(\theta_i,\phi_i) = \sum l_i \tag{14}$$

Where, $\theta_i = \theta(x^t{}_i\beta) = \theta(\eta_i)$

The MLE of β are obtained by maximizing the log- likelihood functions l (β,ϕ). where θ is known and monotone function, and then the likelihood function of GLM depends on β only through linear predictor, η. The MLE of β are the solutions*- of the simultaneous equations of

$$\frac{\partial l(\beta,\phi)}{\partial\beta} = \sum \frac{\partial l_i}{\partial\beta_i} = 0 \tag{15}$$

For detail properties and statistical inference, including the maximum likelihood estimation of the parameters for the Poisson, negative binomial, ZIP, ZINB or hurdle models, refer to [18, 13 and14]

Assessing Model Adequacy

Likelihood Ratio test

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

$$G^2 = -2\ln\left[\frac{L_O}{L_1}\right] = -2\{\ln L_O - \ln L_1\} \sim \chi_p{}^2 \tag{16}$$

Where, L_O and L_1 are the maximized log-likelihood of models under the null and alternative a hypothesis, respectively. This statistic is called the likelihood-ratio test statistic. Comparing k predictor, p is the number of parameters which is the difference between the two likelihood functions and x_p^2 is a chi-square distribution with p degree of freedom. If the test statistics exceeds the critical value, the null hypothesis is rejected. That means the overall model is significant. In this study, to compare Poisson and NB regression models and also ZIP with ZINB regression models, we used significance of dispersion parameter and likelihood ratio (LR) test as criterions. The statistic of likelihood ratio test is given by the following equation:

$$LRT_\alpha = -2(LL_1 - LL_2) \tag{17}$$

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

Akaike Information Criteria (AIC)
In statistical literatures, based on several likelihood measures, one can compare several models performance. One of the most regularly used measures is AIC. For comparison of non-nested models based on maximum likelihood, several authors beginning with [1] have proposed model selection criteria based on the fitted log-likelihood function. The AIC penalized a model with larger number of parameters, and is defined as

$$AIC = -\ln L + 2P \tag{18}$$

Where ln L is the maximized likelihood function of the estimated model and −ln L offers summary information on how much discrepancy exists between the model and the data, where, p is the number of parameters in the model [17]. A relatively small value of AIC is preferred for the fitted model.

Bayesian Information Criteria (BIC)
Unlike, the Akaike information criteria the Bayesian information matrix (BIC) takes into account the size of the data under considered. It is given by

$$BIC = -2\ln L + p\log(n) \tag{19}$$

Where, L is the log likelihood of a model that will compare with the other models, n is the sample size of the data and p is the number of parameters in the model including the intercept. The better model is the one which has the minimum BIC value [17].

Vuong's Test
The Vuong test is a non-nested test that is based on a comparison of the predicted probabilities of two models that do not nest [24]. This test is used for model comparison. Let's define: $m_i = \ln\left(\frac{p_1(y_i / x_i)}{p_2(y_i / x_i)}\right)$ where $p_N(y_i / x_i)$ is the predicted probability of the observed count in case I from model N, then Vuong test statistic test the hypothesis given as:

$$V = \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} m_i\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i - \overline{m})^2}} = \frac{\sqrt{\mathrm{n}}}{\mathrm{s_m}}\overline{\mathrm{m}} \qquad \text{where} \quad \overline{\mathrm{m}} = \frac{1}{n}\sum_{i=1}^{n} m_i$$

where $\overline{\mathrm{m}}$, n and $\mathrm{s_m}$ are mean, sample size and standard deviations respectively.
The hypotheses of the Vuong test are:

$$\mathrm{H_O} : E[m_i] = 0 \quad \text{Vs} \quad \mathrm{H_a} : E[m_i] \neq 0$$

The null hypothesis of the test is that the two models are equivalent. Vuong showed that asymptotically, V has a standard normal distribution. As Vuong notes, the test is directional [24]

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

Test for individual predictors

Consider a null hypothesis $H_O : \beta = 0$, where β is the parameter. Let $SE(\hat{\beta})$ denote the standard error of β, evaluated by substituting the ML estimate for the unknown parameter in the expression for the true standard error. When H_O is true, the test statistics is:

$Z = \frac{\hat{\beta} - \beta_O}{SE(\hat{\beta})}$, has approximately a standard normal distribution.

Equivalently, Z^2 has approximately Chi-squared distribution with $df = 1$. This type of statistic, which uses the standard error evaluated at the ML estimate, is called a Wald test statistic. The Wald test statistic is:

$$Z^2 = \left(\frac{\hat{\beta} - \beta_O}{SE(\hat{\beta})} \right)^2 \quad (20)$$

Under H_0 is true Z^2 have chi-squared distributions with $df = 1$.

3. Discussion of the Results

Children are the human resource bank of every nation. However, they are more vulnerable to diseases and other health risks. Under-five mortality, therefore erodes the potential economic labor force of a country thus plunges the country into human resource crisis and retards development. Therefore, this study has been attempted to identify socioeconomic, demographic, health and environmental related determinants of the number of under-five mortality based on EDHS, 2016 dataset.

According to the results, Mother's educational level has a negative relationship on under-five mortality that is mortality rate decreases with increase in Mother's education level. Thus, the odds of having the opportunity of experiencing under-five deaths for primary and secondary & above educational level were decreased by 54.58% and 72.72%, respectively compared with not educated mothers holding other variables in the model. This result in lined with the previous study that, the higher the level of Mother's educational level, the lower child mortality [11, 12]

In this study the risk of under-five death associated with multiple births is very high relative to single births and this study is similar to the previous studies that birth type to be linked with under-five child death as multiple births is associated with a higher risk of child mortality [11, 12, and 29].

According to the model selection criteria (value of AIC and BIC), the hurdle negative binomial regression model is the most appropriate to fit this data (table1).

From the results, it was found that under-five children who live in Addis Ababa region are found to have a lower risk of under five-child mortality per mother and are 0.3884 times lower than Tigray (reference category). This could be as a result of higher health facilities and living standards in Addis Ababa, compared to other regions. This study is similar with the previous studies that Addis Ababa had a lower hazard (risk) of death (p value = 0.048) [2]. The result is also in lined with [7] by using Ethiopia Demographic and Health Survey (EDHS) 2011 data poison and negative binomial models were used for data analysis and there were regional variations in under-five mortality per mother.

The finding of the study on the variable contraceptive use has negative relationship (-0.50352) on under- five mortality revealed that the death of under-five children from mothers using contraceptive is significantly less than children from non-contraceptive methods using mothers. The result is in accordance with [4, 12].

The study also revealed that Birth order is an important variable that affects the number of under-five mortality. It has a positive effect. Amazingly, as Birth order increases the under-five mortality also increases and this result is consistent with the literature reviewed and contribution from different studies on birth order [5, 29].

From the results, it was found that under-five children who live in Urban are found to have a lower risk of under five-child mortality per mother and are 0.2816 times lower than Rural (reference category). This result is in line with [2] stated under-five children who lived in rural areas had a higher hazard (risk) of death compared to those living in urban areas. In addition, under-five children who lived in rural areas had 18% (p value = 0.01) more hazard (risk) of death than those living in urban areas. This result is also accordance with [5] significantly higher rate of under-five mortality was encountered among rural *kebeles* than in urban *kebeles*.

Table1: Model selection criteria for the count regression models

Model	Poisson	NB	ZIP	ZINB	HP	HNB
AIC	29372.38	27239.76	26730.13	26631.04	26725.53	26620.95
BIC	29557.09	27422.47	26864.83	26763.71	26860.23	26753.65
Log likelihood	-14661.19	-13593.88	-13315.06	-13264.52	-13312.76	-13259.47

Table 2: HNB model results for under-five death of selected independent variables.

Count part	Estimate	S.E	Z-value	P-value	IRR	95% CI of IRR	
						Lower	Upper
Intercept	0.4021	0.0722	5.571	0.0001	1.4950	1.2978	1.7222
Region (Tigray, ref.)							
Afar	0.1484	0.0686	2.161	0.0307	1.1600	1.0139	1.3271
Amahara	0.1510	0.0682	2.212	0.0269	1.1630	1.0173	1.3294
Oromia	0.1184	0.0672	1.762	0.0781	1.1257	0.9868	1.2841
Somali	0.0942	0.0673	1.400	0.1614	1.0988	0.9631	1.2536
Benishangul	0.3613	0.0681	5.307	0.0001	1.4352	1.2558	1.6400
SNNPR	0.1498	0.0666	2.249	0.0245	1.1616	1.0194	1.3236
Gambela	-0.2430	0.0937	-2.593	0.0095	0.7843	0.6527	0.9424
Harari	-0.1751	0.1016	-1.724	0.0847	0.8394	0.6878	1.0242
Addis Ababa	-0.3884	0.1965	-1.976	0.0481	0.6781	0.4613	0.9968
Dire Dawa	0.2501	0.0802	3.117	0.0018	1.2841	1.0973	1.5029
Birth type (Single, ref.)							
Multiple	0.3614	0.0477	7.568	0.0001	1.4353	1.3071	1.5762
Breastfeeding (No, ref.)							
Yes	-0.1327	0.0335	-3.962	0.0001	0.8757	0.8200	0.9351
Source of drinking water (piped, ref.)							
Other water	0.0471	0.0400	1.176	0.2397	1.0482	0.9691	1.1337
Mother's occupation(NO, ref)							
Yes	-0.0573	0.0336	-1.708	0.0876	0.9443	0.8841	1.0085
Mother's education (No education, ref.)							
Primary	-0.4416	0.0520	-8.490	0.0001	0.6430	0.5807	0.7120
Secondary $ above	-0.7693	0.1299	-5.924	0.0001	0.4633	0.3592	0.5976
Contraceptive use (Not, ref.)							
Using	-0.2860	0.0460	-6.214	0.0001	0.7513	0.6865	0.8222
Birth order (first, ref.)							
2-5	0.4249	0.0311	13.664	0.0001	1.5294	1.4389	1.6255
6 & above	0.7446	0.4385	1.698	0.0895	2.1056	0.8915	4.9729
Log(theta)	1.7974	0.1352	13.289	0.0001	6.0339		
Zero-inflation part					OR	95% CI of OR	
						Lower	Upper
Intercept	-0.3283	0.0948	-3.463	0.0005	0.7201	0.5980	0.8672
Region (Tigray, ref.)							
Afar	0.3483	0.0966	3.606	0.0003	1.4167	1.1723	1.7119
Amahara	0.4199	0.0905	4.642	0.0001	1.5218	1.2745	1.8172
Oromia	0.1653	0.0884	1.871	0.0614	1.1797	0.9921	1.4028
Somali	0.0855	0.0937	0.912	0.3616	1.0893	0.9065	1.3088
Benishangul	0.3909	0.0978	3.997	0.0001	1.4783	1.2204	1.7907
SNNPR	0.4669	0.0890	5.248	0.0001	1.5950	1.3398	1.8987
Gambela	0.1755	0.1060	1.656	0.0978	1.1918	0.9683	1.4671
Harari	0.0227	0.1170	0.194	0.8464	1.0229	0.8133	1.2865
Addis Ababa	-0.5123	0.1519	-3.373	0.0007	0.5991	0.4449	0.8069
Dire Dawa	0.4342	0.1106	3.924	0.0001	1.5437	1.2427	1.9174
Birth type (Single, ref.)							
Multiple	2.2782	0.1515	15.041	0.0001	9.7591	7.2522	13.133
Breastfeeding (No, ref.)							
Yes	-0.7182	0.0450	15.964	0.0001	0.4876	0.4465	0.5326
Mother's occupation (No, ref.)							
Yes	0.1327	0.0473	2.807	0.0050	1.1419	1.0408	1.2528
Mother's educational level (No, ref.)							
Primary	-0.7892	0.0557	-14.177	0.0001	0.4542	0.4073	0.5066
Secondary & above	-1.2989	0.0994	-13.066	0.0001	0.2728	0.2245	0.3315
Contraceptive use (Not, ref.)							
Using	-0.5035	0.0545	-9.234	0.0001	0.6044	0.5431	0.6726
Birth type (first, ref.)							
2-5	0.9238	0.0574	16.100	0.0001	2.5188	2.2508	2.8185
6 & above	10.814	195.25	0.055	0.9558	49711	0	7.77e+

Conclusions

The purpose of this study was to identify, socioeconomic, demographic, health and environmental related determinants of number of under-five mortality per mother in Ethiopia. The descriptive results showed that 62.5% of mothers have no experienced under-five death and only 0.7% of them lost at least 7 of their under- five children.

From the exploratory results we could identify that there was an excess zeros and high variability in the non-zero values. The variance of the number of under-five death is larger than its mean, indicating that there is a possibility of over dispersion. In addition to this, the over dispersion parameter is significantly different from zero in NB, ZINB as well as HNB models. This implies that the standard Poisson model is not an adequate model to fit the under-five death per mother.

In this study count regression model is used. There are an excess number of zeros in the dataset. From the results of count regression models, Hurdle negative binomial regression model is better fitted the data which is characterized by excess zeros and high variability in the non-zero outcome than any other models and therefore, it is selected as the best parsimonious model to predict the number of under-five death in Ethiopia. For selected HNB model, the truncated negative binomial part, the explanatory variables like region, mother's education level, contraceptive use, birth order, Breastfeeding and Birth type are statistically significant factors while for logistic part, Using breastfeeding method, Place of residence, Region, Sex of child, mother's occupation, Mother's educational level, Contraceptive use, Wealth index, Marital status, Birth type and Birth order are statistically significant factor on the number of under-five death per mother in Ethiopia.

Recommendations

Based on the findings of the study, we forward the following possible recommendations:

The educational level of mothers plays an important role in child survival, efforts are needed to extend educational programmers aimed at educating mothers. Health programs need to focus on supporting women with little or no education in order to reduce under- five mortality.

Policies and programs aimed at addressing regional differences in under-five mortality must be formulated and their implementation must be vigorously pursued. To achieve this, while under-five mortality reduction measures which have worked to some extent in the Affar, Amhara, Benshangul-Gumuz, SNNP and Dire Dawa must be strengthened to achieve more results in the region, these measures could be extrapolated and applied in the remaining regions of the country.

The concerned body should work closely with both the private sector and civil society to teach households to have sufficient knowledge and awareness on under five mortality and mechanisms of reduction and moreover to make children very well.

Reference

1. Akaike H. Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika. 1973; 60(2):255-265.
2. Ayele DG, Zewotir TT, Mwambi H. Survival analysis of under-five mortality using Cox and frailty models in Ethiopia. Journal of Health, Population and Nutrition. 2017; 36(1):25.
3. Baksh MF, Böhning D, Lerdsuwansri R. An extension of an over-dispersion test for count data. Computational Statistics & Data Analysis. 2011; 55(1):466-474.
4. Bedada D. Determinant of Under-Five Child Mortality in Ethiopia. American Journal of Statistics and Probability. 2017; 2(2):12-18.
5. Bedane AS, Asena TF, Shamenna AT, Abshoko AD. Variations in Under five Child Mortality among Regional States of Ethiopia: A Multi-level Modelling Approach. health, 2016; 5:6.
6. Bereka SG, Habtewold FG, Nebi TD. Under-Five Mortality of Children and its Determinants in Ethiopian Somali Regional State, Eastern Ethiopia. Health Science Journal, 2017; 11(3).
7. Berhie KA, Yirtaw TG. Statistical Analysis on the Determinants of Under Five Mortality in Ethiopia. American Journal of Theoretical and Applied Statistics. 2017; 6(1):10-21.
8. Cameron AC, Trivedi PK. Econometric models based on count data. Comparisons and applications of some estimators and tests. Journal of applied econometrics. 1986; 1(1):29-53.
9. Ethiopia Demographic and Health Survey Addis Ababa, Ethiopia, 2016.
10. Gardner W, Mulvey EP, Shaw EC. Regression analyses of counts and rates: Poisson, over-dispersed Poisson, and negative binomial models. Psychological bulletin. 1995; 118(3):392.
11. Gebretsadik S, Gabreyohannes E. Determinants of under-five mortality in high mortality regions of Ethiopia: an analysis of the 2011 Ethiopia demographic and health survey data. International Journal of Population Research, 2016.
12. Getachew Y, Bekele S. Survival Analysis of Under-Five Mortality of Children and its Associated Risk Factors in Ethiopia. J Biosens Bioelectron. 2016; 7(213):2.
13. Gurmu S. Generalized hurdle count data regression models. Economics Letters. 1998; 58(3):263-268.
14. Hilbe JM. Negative binomial regression. Cambridge University Press, 2011.
15. Hoffmann JP. Generalized linear models: An applied approach. Pearson College Division, 2004.
16. Jansakul N, Hinde JP. Score tests for zero-inflated Poisson models. Computational statistics & data analysis. 2002; 40(1):75-96.
17. Konishi S, Kitagawa G. Information criteria and statistical modeling. Springer Science & Business Media, 2008.
18. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics. 1992; 34(1):1-14.
19. Long SJ, Long JS, Freese J. Regression models for categorical dependent variables using Stata. Stata press, 2006.
20. McCullagh P, Nelder. Generalized Linear Models. London, chapman hall, 1989.
21. Mullahy J. Specification and testing of some modified count data models. Journal of econometrics. 1986; 33(3):341-365.
22. Mwalili SM, Lesaffre E, Declerck D. The zero-inflated negative binomial regression model with correction for misclassification: an example in caries research. Statistical methods in medical research. 2008; 17(2):123-139.
23. UNICEF. Levels & Trends in Child Mortality. Estimates developed by the UN Inter agency Group for Child Mortality Estimation, 2018.
24. Vuong QH. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica: Journal of the Econometric Society, 1989, 307-333.

25. WHO. World health statistics 2016: monitoring health for the SDGs sustainable development goals. World Health Organization, 2016.
26. WHO. UN Inter-agency Group for Child Mortality Estimates, 2017.
27. WHO. UNICEF, World Bank, Levels & Trends in Child Mortality. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation, 2018.
28. You D, Hug L, Ejdemyr S, Idele P, Hogan D, Mathers C, *et al*. Global, regional, and national levels and trends in under-5 mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. The Lancet. 2015; 386(10010):2275-2286.
29. Zike DT, Fenta HM, Workie DL, Swain PK. Determinants of Under-Five Mortality in Ethiopia: An Application of Cox Proportional Hazard and Frailty Models. Turkiye Klinikleri Journal of Biostatistics, 2018; 10(2).
30. Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. GLM and GAM for count data. In Mixed effects models and extensions in ecology with R. Springer, New York, NY, 2009, 209-243.