Count Models for Identifying Factors Associated with Child Mortality in Ethiopia

Endale Alemayehu Ali (endalestat@gmail.com)
Ambo University

Tsigereda Tilahun
Ambo University

Research

Keywords: Child mortality, Counting Models, Zero-inflated negative binomial regression mod

Posted Date: February 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-222761/v1

Count Models for Identifying Factors Associated with Child Mortality in Ethiopia

Endale Alemayehu[1*], Tsigereda Tilahun[2]

Email: endalestat@gmail.com, tisgeti@gmail.com

[1*, 2] Department of Statistics, College of Natural and computational Science, Ambo University

Abstract

Introduction: The under-five mortality rate, often known by its acronym U5MR or simply as the child mortality rate, indicates the probability of dying between births exactly five years of age, expressed per 1,000 live births. In comparison, the probability of dying after the first month and before reaching age 1 was 12 per 1,000, the probability of dying after age 1 and before age 5 was 10 per 1,000, and the probability of dying after age 5 and before age 15 was 7 per 1,000.

Objectives: The study was aimed to determine the major factors of child mortality in Ethiopia using different counting models. In detail the study has the objective of identifying the risk factors of child mortality in Ethiopia and also to prioritize the best counting models that fit the data well.

Methods: The Ethiopian demographic and health survey of 2016 was used for this study. About 10641 women aged between 15-49 were included in the survey. To analyze the data, counting models like the Poisson regression model, negative binomial model, zero-inflated regression models, and zero-inflated negative binomial regression model were applicable.

Results: The results of the study indicated that of the total 10641 women respondents, 7576 (71.2%) have not faced the problem of child mortality. Thus, this result has the clue that the count models, especially the models that can handle the dispersion may be applicable. The average rate of child mortality is less than the variance of child mortality and this indicated that there is an over-dispersion of the data. Of all the candidate models, a zero-inflated negative binomial regression model was found to be the best model since it has a minimum AIC(15517). The coefficient table of the best model indicated that of child mortality for the women from rural residence is 1.2532 greater than those from urban with a 95% confidence interval (0.0905, 0.3610).

Conclusion: The model comparison technique is indicated that the zero-inflated negative binomial regression models were the best mode that fit the data well. Under this model, the residency of women, birth order, Preceding Birth Interval, Size of a child at birth (smaller than average), and number of household members are significant variables in determining the status of child mortality in Ethiopia

Key words: Child mortality, Counting Models, Zero-inflated negative binomial regression mod

1. Backgrounds

Child mortality is among the factors that can be associated with the well-being of a population and taken as one of the development indicators of health and socioeconomic status and also indicates a life quality of a given population, as measured by life expectancy. Hence its indication is a very important for evaluation and public health strategy. Thus it is an area that many researchers focus and that has attracted the attention of policy-makers and program implementers worldwide (Unicef, 2018).

Globally, the number of years that a newborn is expected to live, given the prevailing and risks of mortality, increased by 24 years from 1950 to 2015, which represents a rise of about 3.6 years per decade over the past 65 years.Children's problems in developing countries are quite different from those in industrial areas. These are many more children died they might compare 40% of the population. Usually one half of all deaths take places in the age group below five years in which the major health problems of the countries are concentrated. More than 97% of all deaths takes place below age five year in less developed countries (Shiferaw, 2018).

Among children and young adolescents, the risk of dying was highest in the first month of life at an average rate of 18 deaths per 1,000 live births (including children under age 5, children aged globally in 2017. In comparison, the probability of dying after the first month and before reaching age 1 was 12 per 1,000, the probability of dying after age 1 and before age 5 was 10 per 1,000, and the probability of dying after age 5 and before age 15 was 7 per 1,000 (Unicef, 2018).

The risk of a child dying before completing five age is still the highest in the World Health Organization(WHO)African countries (81per 1000 live births), about 7 times higher than in theWHO European region (11 per under-five mortality rates per 1000 live births). Low-income countries reported 76 deaths per 1000 live births, about 11 times the average rate in high- income countries (7 deaths per 1000 live births). All 16countries with anunder-fivemortality rate above 100deathsper 1,000 live births are in sub-SaharanAfrica (Gebreweld et al., 2018).
There are huge differences in child mortality among low and middle income countries and the industrial world with Sub-Saharan Africa and South East Asia carrying the highest burden of under-five mortality. The highest rates of child mortality are still in Sub-Saharan Africa-where 1 in

8 children dies before age 5, more than 17 times the average for developed regions (1 in 143)-and Southern Asia (1 in 15). Ethiopia being one these countries under five mortality still The study design was a retrospective study remain high, 88/1000 live birth. In Ethiopia, mainly in the Somali region under-five mortality was well above the national average (Bereka et al., 2017; Susuman, 2014).

The studies indicated that after a decade, out of the 187 countries, only nineteen countries- all in Africa- had an infant mortality rate of above ten percent including Ethiopia. Ethiopia is one of those countries infant and child mortality is also high due to among the factors putting infants at risk of illness and death in Ethiopia are short birth intervals, high birth order, low birth weight, the age and educational attainment of the mother, tetanus and other infections, congenital factors, and being part of certain religious group. Maternal education has been observed to be strong predicator of the infant and child mortality in developing countries. The education of the mother is emerged as one of the strongest predicators of child mortality rates through other factors like women's income, working status, standard of living index, and place of residence (Shiferaw, 2018; Dejene & Girma, 2015).

The current levels of under-five mortality still high and seem the objective of reducing the problems in the sustainable development goal may not go to the extent it has to be. Even though the trend of child mortality indicated that there were slight change of the reduction, but still the problem was too crucial. Like other developing countries, significant differentials in mortality levels were observed among rural and urban residents of Ethiopia.
Thus, considering the seriousness and gaps seen in this regards, this study attempt to answer the following basic research questions

The general objective of the study is to determine the major factors of child mortality in Ethiopia using different counting models and identifying the better model that fit the data

2. Methodology

2.1 Source of Data

This study was conducted based on the database that has been compiled as part of the 2016 Ethiopia Demographic and Health Survey (EDHS). It is the fourth Demographic and Health

Survey conducted in Ethiopia. The 2016 EDHS sample was stratified and selected in two stages. Each region was stratified into urban and rural areas. In the first stage, a total of 645 enumeration areas (EAs) were selected with probability proportional to EA size, of which 202 in urban areas and 443 in rural areas. An EA is a geographic area covering on average 181 households. In the second stage of selection, a fixed number of 28 households per enumeration area were selected from the newly created list of household listing using systematic sampling. Nationally 10,641 women of age 15-49 were included in the survey.

2.2 Variables in the study

2.2.1 Dependent variable

The dependent variable of the study was number of Mortality of Children per women counted as 0, 1, 2,..

2.2.2 Independent Variable

The Quantitative variables are Age of mother's at first birth (continuous predictor), Birth order number (count), Preceding Birth Interval (count). The covariate predictor variable given in table 1

Table 1: Description of covariates

Variable Name	Variable description	Category
Residence	Type of place of residence	1=Urban 2=Rural
Education	The highest educational level the woman attained	0=No Education 1=Primary 2=Secondary 3= Higher
Maritalstat	Current marital status of mother's	0=Not Married 1=Married
Size	Size of child at birth	0=Smaller than average 1=Average 2=Larger than average

Variable Name	Variable description	Category
Source	Source of drinking water	1= Unprotected 2= Protected
Month	Month of breast feeding	0=< 6 months 1=>6 months

2.3 Count data models

One of the crucial questions in statistical analysis of count data is how to formulate an adequate probability model to describe observed variation of counts. Because of the restrictive nature of equi dispersion assumption in standard Poisson model, researchers have developed techniques and tests that allow detecting the over dispersion (or under-dispersion) in the population. Recently, Zero-inflated models have been developed to take into account the excess of zeroes in the data. The Zero-inflated model can be seen as a finite mixture model where one distribution is considered as a degenerate process with a unit point mass at zero (Lambert, 1992).

2.3.1 Poisson regression model (PR)

Consider we have n independent random variables denoted by $Y1, Y2, \dots, Yn$. Where y_i denote child death for ith mother in her life time. Poisson regression model for count data assumes that each observed count yi is drawn from a Poisson distribution with rate of occurrence of an event, $parameter \;\; \lambda i \; i = 1, 2, \dots, n.$ Let Xi denote a vector of covariates in the study for the ith mother.

Then the Poisson equation of the model with rate parameter λi is given by:

$$p(Y_i = y_i) = \frac{e^{\lambda_i}}{y_i!} {\lambda_i}^{y_i}, \;\; \lambda_i > 0 \;\; and \;\; y_i \;\; = 1, 2, 3, \dots$$

[1]

The mean of the response variable λ_i is related with the linear predictor through the so called link function (Nelder and Wedderbum, 1972).

The relationship between Yi and ith row vector of X, X_i, linked by $g(\lambda_i)$, is the canonical link function given by:

$$E(Y_i) = \lambda_i = e^{X_i'\beta} \quad [2]$$

2.3.2 Negative binomial regression model (NBR)

The main problem in the use of standard Poisson regression model, where there a more dispersed observations or if the data have excess zeros, another way of modeling such over-dispersed count data is a negative binomial (NB) distribution which can arise as a gamma mixture of Poisson distributions.

Let $Y1, Y2, \ldots, Yn$ be a set of n independent random variables one parameterization of its probability density function is, then the negative binomial model given as:

$$P(Y_i = y_i) = \frac{\Gamma(y_i+\frac{1}{\alpha})}{\Gamma(y_i+1)\Gamma(\frac{1}{\alpha})}\left(\frac{\alpha\lambda_i}{1+\alpha\lambda_i}\right)^{y_i}\left(\frac{1}{1+\alpha\lambda_i}\right)^{1/\alpha} \quad , \; y_i = 0, 1, 2, \ldots and \quad \alpha \geq 0$$

[3]

Covariates can be introduced into a regression model based on the NB distribution via the relationship:

$$ln(\lambda_i) = \beta_0 + \sum_{j=1}^{k} x_j \beta_j \quad [4]$$

Where xj is the jth covariate, k is the number of covariates in the model and βj is the jth regression parameter (Lawless 1987; Cameron & Privedi 1998).

2.3.3 Zero-inflated regression models

One characteristic of the Poisson distribution is that the mean of the distribution is equal to the variance; however when there are excess zeros, probability of zero in the standard model will be less than the expected. The problem of standard models in under predicting zeros and over

predicting ones is very common and sometimes this problem can be very severe when there are a lot of zeros in the distribution. In such cases, Zero inflated Poisson (ZIP) and Zero inflated negative binomial (ZINB) models can be used to account for excess zeros. The zero values in the ZIP model can be viewed as comprising two parts. One portion of the zero counts arises from the inflated part of the distribution and the other portion comes from what would be expected given a Poisson distribution with parameter λ (Natalia & Gerstel, 2004).

Zero-inflated negative binomial regression model (ZINB): The main difference between ZIP and ZINB model is that the Poisson distribution for the count data is replaced by the negative binomial distribution.

The ZINB model is a special case of a two-class finite mixture model like the ZIP model with mean $E(Yi) = \lambda i\ (1 - \Phi i)$ and variance $Var\ \ (Yi) = \lambda i\ (1 - \Phi i)\ \ (1 + \alpha\lambda i + \Phi i\lambda i)$, where the parameters $\lambda i\ and\ \ \Phi i$ depend on the covariates and $\alpha \geq 0$ is a scalar. Thus we have over-dispersion whenever either Φi or α is greater than 0. Thus, the equation in (10) reduces to NB when $\Phi i = 0$ and to the ZIP when $\alpha = 0$.

Test for over dispersion parameter: The negative binomial regression model reduces to the Poisson regression model when the over dispersion parameter is not significantly different from zero.

This is a test of significance of the over dispersion parameter α. The presence of the over dispersion parameter α in the NB regression model is justified when the null hypothesis H_0: α=0 is rejected. In order to test the above hypothesis a score test statistic is proposed.

Goodness of-fit tests: In this section, several goodness-of-fit measures will be briefly discussed, including the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC).AIC is the most common means of identifying the model which fits well by comparing two or more than two models. It tries to balance the goodness of fit against the complexity of the model (Schwarz, 1978).

3. Results

3.1 Descriptive statistics

The descriptive statistics given in the Table 2 shows the number and percentage of mortality of children. The result shows that 71.2% of the mothers have not faced any child death. Since 71% of the data is zeroes thus zero inflated modal may appropriate.

Mortality of children	Frequency	Percent
0	7576	71.20
1	2021	18.99
2	633	5.95
3	255	2.40
4	104	0.98
5	33	0.31
6	15	0.14
8	1	0.01
9	1	0.01
10	2	0.02
Total	10641	100

Table 2: Number of mortality of children per women in Ethiopia

3.2 Test for over dispersion

In Poisson regression analysis, deviance and Pearson Chi square goodness of fit statistics indicate there was over dispersion (Table 3). Since the Pearson Chi square statistic divided by the degrees-of-freedom is higher than one and the observed value of 1.1, then the mentioned goodness of statistics represents that there was an over dispersion in the data set. Even if the Deviance and Pearson Chi square goodness of fit statistics of 8165.60 and 11118.57 respectively in Negative Binomial regression is dropped considerably, still significant over dispersion exists; because we would like to divide this value by the degrees of freedom to be close to one.

Table 3: Test for Over Dispersion

Models		DF	Value	Value/DF
Deviance	Poisson	10627	8023.4950	0.7550
	Negative Binomial	10627	8165.6039	0.7683
Pearson	Poisson	10627	11436.3971	1.0762
	Negative Binomial	10627	11118.5671	1.0463

3.3 Model selection criteria

As shown in Table 4 ZIP and ZINB regression models were better fitted than Poisson and NB respectively based on their corresponding AIC. It is found that the models with the smallest AIC and BIC were ZIP regression followed by ZINB regression model.

Table 4: Model selection criteria for the regression models

Criterion	P	NB	ZIP	ZINB
AIC	15696	15649	15526	15517
BIC	15698	15758	15729	15728

The smallest AIC and BIC is the best model in here ZINB is the best model then we fit the model with ZINB.

3.4 Parameter estimates and interpretation

Interpretation (zero-inflated model)

Results in Table 5 provide estimates of the effect of some selected characteristics of mothers on the mortality of their children. These coefficients can be interpreted the same way as regular negative binomial coefficients. The factors place of residence, number of house hold members, preceding birth interval, size of child at birth and birth order number are statistically significant determinant of child mortality in Ethiopia. Age of mother at time of first birth, Month of breast feeding and women's education level are insignificant factors of child mortality in Ethiopia.

Result in Table 5 indicates children living in the rural areas had an increased risk to death compared to those living in the urban areas. The risk of child mortality was about 1.2532 times higher for child whose mother resides in rural areas compared to their urban counterparts (OR: 1.2532 95%CI: 0.0905-0.3610).

The risk of child mortality for those children whose size of child at birth larger than average is 87% higher compared to those who size of child at birth smaller than average. (OR: 0.8707 95% CI:-0.2224-(-0.0543). There is no significant difference in child mortality among those size of child at birth average and smaller than average.

As preceding birth interval increases by one, the expected number of zero death of children per women decreases by 0.0009. As birth order number increases by one, the expected number of zero death of children per women increases by 0.0101.

As a number of household members increases by one, the expected number of zero death of children per women decrease by 0.2664.

Parameter		DF	Estimate	Standard Error	Exp(B)	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.7967	0.2320	0.4508	-1.2515	-0.3419	11.79	0.0006
Age at first birth		1	-0.0015	0.0051	0.9985	-0.0114	0.0085	0.08	0.7746
Birth order number		1	0.3648	0.0101	1.4402	0.3451	0.3845	1316.89	<.0001
Preceding Birth Interval		1	-0.0067	0.0009	0.9933	-0.0085	-0.0049	52.23	<.0001
Types of Residence (urban)	Rural	1	0.2257	0.0690	1.2532	0.0905	0.3610	10.70	0.0011
Education (No education)	Higher	1	-0.4334	0.3227	0.6483	-1.0659	0.1991	1.80	0.1793
	Primary	1	-0.0929	0.0526	0.9113	-0.1960	0.0102	3.12	0.0774
	Secondary	1	0.2006	0.1235	1.2221	-0.0414	0.4426	2.64	0.1043
Source of water (Unprotected)	Protected	1	0.0107	0.0340	1.0107	-0.0558	0.0773	0.10	0.7518
Marital status of mothers(Notmarried)	Married	1	0.1023	0.1818	1.1077	-0.2540	0.4586	0.32	0.5736
Size of child at birth (smaller than average)	Average	1	-0.0583	0.0398	0.9434	-0.1362	0.0197	2.15	0.1429
	Larger than Average	1	-0.1384	0.0429	0.8707	-0.2224	-0.0543	10.42	0.0012
Month of breast feeding (less than 6	greater than 6	1	0.0648	0.0460	1.0669	-0.0253	0.1549	1.99	0.1585

months)	months								
Number of house hold members		1	-0.2664	0.0099	0.7661	-0.2858	-0.2469	719.66	<.0001
Dispersion		1	0.0750	0.0243		0.0397	0.1416		

Table 5: Parameter estimates of the ZINB Model

3.5 Discussion

The descriptive results of the study indicated that of the total 10641 mother's, most of them 7576 (71.2%) of the mothers have not faced any child death. Many other previous studies also revealed that the mortality of child.

Children living in the rural areas had an increased risk to death compared to those living in the urban areas. The risk of child mortality was about 1.2532 times higher for child whose mother resides in rural areas compared to their urban counterparts. Theoretically, all things being equal, living in urban areas should be associated with a higher standard of living, better sanitation and better health facilities, among other things. Place of residence was found to have an association with child mortality. This finding is in agreement with studies by which shows a significant association between place of residence and child mortality. (Berhanu, 2018)

The result of this study shows that household members has a negative association with child mortality. As a number of household members increases by one, the expected number of zero death of children per women decrease by 0.2664. Children born in a family size of child at birth larger than average have a 0.8707 times higher risk of death as compared with those who size of child at birth smaller than average. In contrast a study conducted in Ethiopia by Desta obtained a negative relationship between family size and child mortality as the family size increased, child mortality becomes lower (Desta, 2011).

The result of this study also showed that preceding birth interval has a negative association with child mortality. As preceding birth interval increases by one, the expected number of zero death of children per women decrease by 0.0009. A study in Ethiopia showed that children born after 18–23, 24–35, 36–47, and more than 47 months of the preceding birth intervals have lower risk of

child mortality by 49, 78, 79, and 89 percent, relative to children born after less than 18 months, respectively (Desta 2011).

The result also showed that the birth order number has a positive association with \child mortality. As birth order number increases by one, the expected number of zero death of children per women increase by 0.0101. A study in Ethiopia showed higher birth orders (>4) have the highest mortality risk. Child with these characteristics are 2.067 times more likely to die in age less than 5 relative to the reference group births of order one. (Solomon et al, 2017).

The study was aimed to determine the major factors of child mortality in Ethiopia using different counting models. The result shows that 71.2% of the mothers have not faced any child death. Since 71% of the data is zeroes thus zero inflated model may appropriate. The mean of the mortality of children was 0.45 and the variance was 0.76 this indicate the variance is greater than the mean value then it is over dispersion.

In Poisson regression analysis, we can also check over dispersion by using deviance and Pearson Chi square goodness of fit statistics indicate there was over dispersion. Since the Pearson Chi square statistic divided by the degrees-of-freedom is higher than one and the observed value of 1.1, then the mentioned goodness of statistics represents that there was an over dispersion in the data set.

In this study, after fitting four different count models it was found that ZIP and ZINB regression models were better fitted than Poisson and NB respectively based on their corresponding AIC. It is found that the models with the smallest AIC and BIC were ZIP regression followed by ZINB regression model.

According to ZINB the factors place of residence, number of house hold members, preceding birth interval, size of child at birth and birth order number are statistically significant determinants of child mortality in Ethiopia. Age of mother at time of first birth, Month of breast feeding and women's education level are insignificant factors of child mortality in Ethiopia.

1. Declarations

Ethics approval and consent to participate

Not applicable

Consent of Publication

This study is our original work and is not published elsewhere before

Availability of data and materials

The data is available online:

https://dhsprogram.com/data/dataset/Ethiopia_Standard-DHS_2016.cfm?flag=0

Competing interests

The authors declare that they have no competing interest

Funding

Not applicable

Authors' contributions

This study was designed and compiled by both EA and TT. The development of the basic research questions, identifying the problems and selecting appropriate statistical models have been done with the collaboration of the authors

Consent for Publication

Not Applicable

Acknowledgement

Not Applicable

Abbreviations

AIC: Akaike Information Criteria, BIC: Bayesian Information Criteria, CI: Confidence Interval, DF: Degree of Freedom, EAs: Enumeration Areas, EDHS: Ethiopian Demographic and Health Survey, NB: Negative Binomial, ZINB: Zero Inflated Negative Binomial, ZIP: Zero Inflated Poisson

References

Bereka, S. G., Habtewold, F. G., & Nebi, T. D. (2017). *iMedPub Journals Under-Five Mortality of Children and its Determinants in Ethiopian Somali Regional State , Eastern Ethiopia* 1–8. https://doi.org/10.21767/1791-809X.1000507

Dejene, T., & Girma, E. (2015). *Social determinants of under-five mortality in Ethiopia : Event history analysis using evidence from Ethiopian Demographic and Health Survey (EDHS) *. 5*(5), 879–884.

Developed, E., Group, I., Mortality, C., & Estimation, M. (2018). *Child Mortality 2018*.

Gebreweld, F. H., Kifle, M. M., Gebremicheal, F. E., Simel, L. L., Gezae, M. M., Ghebreyesus, S. S., Mengsteab, Y. T., & Wahd, N. G. (2018). Factors influencing adherence to tuberculosis treatment in Asmara, Eritrea: a qualitative study. *Journal of Health, Population and Nutrition*, *37*(1), 1. https://doi.org/10.1186/s41043-017-0132-y

Shiferaw, Y. A. (2018). *DETERMINANT OF INFANT AND CHILD*. *May 2012*. https://doi.org/10.2139/ssrn.2188355

Susuman, A. S. (2014). *Child Mortality Rate in Ethiopia*.

Lambert D (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 34: 1-14.

EDHS: Ethiopian demographic and health survey report. 2016.

Nelder JA, Wedderburn RWM (1972) Generalized linear models. J R Stat Soc 135: 370-384.

Lambert D (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 34: 1-14.

Lawless JF (1987) Negative binomial and mixed Poisson regression. Can J Stat 15: 209-225.

Cameron AC, Trivedi PK (1998) Regression analysis of count data. Cambridge University Press

Natalia S, Gerstel N (2004) Explaining the gender gap in help to parents: The importance of employment. J Marriage Fam 66: 431-451.

Long JS (1997) Regression models for categorical and limited dependent variables. Thousand Oaks, CA: Sage.

Desta M (2011) Infant and child mortality in Ethiopia: The role of socioeconomic, demographic and biological factors in the previous five years period of 2000 and 2005, Lund university, Lund, Sweden.

Solomon Gebretsadik Bereka, Fekade Getabil Habtewold and Tewodros Desalegn Nebi (2017) Under-Five Mortality of Children and its Determinants in Ethiopian Somali Regional State, Eastern Ethiopia.