2019-08-18

Determinant of Fertility Status among Married Women in Ethiopia: Application of Multilevel Count Regression Model

Abebech Fentie

BAHIR DAR UNIVERSITY

SCIENCE COLLEGE

DEPARTMENT OF STATISTICS

DETERMINANT OF FERTILITY STATUS AMONG MARRIED WOMEN IN ETHIOPIA: APPLICATION OF MULTILEVEL COUNT REGRESSION MODEL.

BY

ABEBECH FENTIE

A RESEARCH PROPOSAL SUBMITTED TO DEPARTMENT OF STATISTICS COLLEGE SCIENCE BAHIR DAR UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN BIOSTATISTICS

February, 2019

BAHIR DAR, ETHIOPA

Declaration

I, the undersigned, declare that this thesis is my original work, has not been presented for degree of masters in any other University and that all sources of materials used for this thesis has been duly acknowledged. This thesis has been submitted in partial fulfillment of the requirements for the Degree of Masters of Science in Statistics (Biostatistics) at Bahir Dar University. Therefore, we recommend it to be accepted as fulfilling the thesis requirements.

Name Abebech Fentie

Signature: ___________

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

Mr. Demeke Lakew (Assistant Professor) ___________ ______________

Name of advisor Signature Date

Approval Sheet

We, the undersigned, members of the Board of examiners of the final open defense, the thesis prepared by Abebech Fentie entitled: **"Determinant of fertility status among married women in Ethiopia: application of multilevel count regression model''** and submitted in partial fulfillment of the requirements for the Degree of Masters of Science in Statistics (Biostatistics) complies with the regulations of the University and examined the candidate ,meets the accepted standards with respect to originality and quality.

Approved by the Board of Examiners:

______________________	______________	________________
Chairperson	signature	date
______________________	______________	________________
Internal Examiner	signature	date
______________________	______________	________________
External Examiner	signature	date

Acknowledgments

First of all, I would like to thanks **God** and Science College in Bihar Dar University for providing me in my life and getting chance to pursue my graduate study.

I would like to gratefully and sincerely thank my thesis advisor, for his consecutive advice and support ideas starting from selecting the study title to the accomplishment of this study.

Furthermore, I would like thanks to the Central Statistical Agency (CSA) of Ethiopia for releasing the data used for this study. In addition I would like to thank my friends and classmates that contributed in diverse ways to my success in this program.

Finally, I would like to express my family, both home and abroad for their continued support, prayer, contributions and bearing with me throughout this program of study.

Abstract

Background: *Fertility is one of the elements in population dynamics that has significant contribution towards changing population size and structure over time in the world. This study investigates determinant factors of fertility among women in Ethiopia.*

Methodology*: The data used for the analysis was obtained from the 2016 Ethiopia Demographic and Health Survey which was implemented by the Central Statistical Agency. The survey collected a total of 15,683 successfully interviewed women aged from 15-49 years out of this 9,602 women were considered in this study. Multilevel Negative binomial analysis was selected to investigate the effect of socioeconomic, demographic, environmental and health related factors on the number of children ever born per woman in Ethiopia.*

Results: *Likelihood ratio test suggested that, the number of ever born children varies across regions and multilevel count regression model was better fit than the single level count regression model. The expected number of total children ever born for using Contraceptive use were lower than 0.9557 times as compared non-using Contraceptive mothers. The expected number of ever born child of farther who has job attachment is higher than 1.05 times that of father who has job attachment.*

Conclusion: *based on the result we can conclude that wealth index, year of education and age of mother at first sex, contraceptive use are negatively associated with total number of ever born children. However, family size, age of mother, age of mother at first birth, number of living child's, father's occupation, region, ideal number of children and number of visits are positively associated with a total number of ever born children were as significant factors. Therefore, the government and other stakeholders should pay attention to the subject and develop intervention to improve on the significant factors identified on the current paper.*

Keywords*: Fertility rate; total children ever born; Ethiopia; ordinary count regression; multilevel regression model.*

Acronyms

AIC	Akaike's information criterion
BIC	Bayesian information criterion
CSA	Central Statistical Agency
EAs	Enumeration Areas
EDHS	Ethiopian Demographic and Health Survey
GEE	Generalized Estimating Equation
GLS	Generalized Least Squares
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimation
NB	Negative Binomial
PHC	Population and Housing Census

Table of Contents

List of Table

List of Figure

CHAPTER ONE

1. Introduction

1.1. Background of the study

Fertility is one of the elements in population dynamics that has significant contribution towards changing population size and structure over time in the world(Champion, 2001). Fertility and future projected population growth are much higher in sub-Saharan Africa than in any other region of the world, and the decline in birth rates, which was already modest, has slowed even further over the past decade(Bongaarts, 2008). About eight percent of the world's population lives in "high-fertility" countries that have experienced only limited fertility decline to date, most of sub-Saharan Africa countries on average has five or more children per woman over her lifetime (UN, 2015b).

Globally, total fertility is 2.5 children per woman in the world, however remains in Africa region with the highest fertility at 4.7 children per woman with particularly high levels in sub-Saharan Africa (5.1 births per woman in 2010-2015) and Europe has the lowest fertility of 1.6 children per woman(UN, 2015a). The world population is estimated to grow from 6.8 billion in 2009 to 9.2 billion in 2050. Three fourths of the world's population lives in developing countries(Alene and Worku, 2009), the average number of births per woman in Sub-Saharan Africa (SSA) is 5.1, which is almost twice that of South Asia (2.8), Latin America and the Caribbean (2.2) (Bank, 2009).

In sub-Saharan Africa, twenty nine countries including Ethiopia are characterized by high fertility rate(Forum, 2012). Most of the sharp rise in population comes from sub-Saharan Africa and in parts of Asia where the number of births will continue to grow well into the 2020's, even if fertility continues to drop(Forum, 2012). These are areas where the protection of adolescents and young women against early or unwanted pregnancy is most inadequate, mortality from unsafe abortion most pronounced, giving birth most hazardous and childhood most difficult to survive(Organization, 2005).

Ethiopia is one of the developing countries with high fertility and rapid population growth rate also it placing the second most populous nation in sub-Saharan Africa country's population in 2016 was estimated around 100 million(EDHS, 2016b),(Forum, 2012).

According to EDHS report Fertility declined only slightly between 2000 and 2005, from 5.5 children per woman to 5.4, and then decreased further to 4.8 children in 2011and finally decreased 4.8 in 2011 to 4.6 in 2016 children per woman implies that a single woman is having more than four children throughout her life(EDHS, 2016b). Despite slight decrement, still high fertility rate of African countries as compared to the world average of 2.5 children per women(Bureau, 2015).

High fertility has adverse effects on the health of children and mother also child schooling. In countries like Ethiopia where the livelihood of about eighty-five percent of the population depends on agricultural practices on small individual holdings, continuous population growth may result in environmental degradation, which ultimately contributes to global warming(Forum, 2012).

Therefore, the aim of this study is investigating the factors that affect number of children ever born per married women in Ethiopia and to identify areas to be focused on in terms of programmatic and policy directions in order to sustain the reduction of fertility of married women. In general this study is fill the gaps that existed in the country like to balance the number of children ever born per mother and resource on the regions, to balance the number of children ever born per mother between regions, etc.

1.2. Statement of the problem

High fertility in Ethiopia remains the dominant factor dictating the future size, growth and composition of the population in the country but Population growth and decline are mainly affected by fertility, in addition fertility is affected by different socio economic and demographic determinants(Adugna, 2014). High fertility hinders achieving national goals such as food sufficiency, universal primary education, improving the accessibility of health care services to the largest possible number in the shortest possible time and employment and housing conditions are remarkably difficult ,and mothers with higher parities have greater risk of death related to pregnancy and child birth and children from large family size have high risk of mortality (UN, 2009, Atsbaha et al., 2016). In order to reduce high fertility rate and control population growth of the country, first we should be clearly identify the factors that influence children ever born per married women.

Also Ethiopian government developed several strategies and plans to reduce total fertility rate from 7.7 children per woman to 4.0 by 2015(Coale and Hoover, 2015),(Nations, 1993)But based on current 2016 EDHS data on average fertility rate is more than four children's per women's(EDHS, 2016a), due to this reason we should focus on the determinant of children ever born per married women classical .

A number of studies conducted by(Priya and Kshatriya, 2016), and (Berlie and Alamerew, 2018); and also by(Awad and Yussof, 2017) using single level classical multiple linear regression and time-serious modeling approach's respectively in the study area on the determinants of number of children ever born among married women. But the drawback of this study the outcome variable number of children ever born is count data not continuous. Since the nature/properties of data and the selected model to analysis the data are misused which generates that miss-conclusion. There are also many studies done by (Atsbaha et al., 2016); (Abebe et al., 2018) on the determinant of fertility status by using binary logistic regression model. However, binary logistic regression model under counts the total number of children ever born since multiple response are collapsed/ aggregated into a two unit to fulfill the requirements of binary logistic regression. Besides, binary logistic regression can't provide sufficient information for studying the pattern of multiple child born that means it merely predicts high /low rather than the number of children ever born. And also this four studies are focused on limited specific areas and descriptive statistical with single level analysis. Therefore, due to this reason to study fertility status of married women use count regression models is more appropriate than multiple linear regression model, time serious regression model and binary logistic regression model of analysis.

Moreover, studies have been done by (Pandey and Kaur, 2015), (Cesar Augusto Oviedo Tejada1, 2017) and (Muluneh, 2016) on fertility status by using count regression model, this study better as compared to other study listed from in the above. But in Ethiopia there have been regional variations in fertility status of married women(EDHS, 2016a), (Eyasu, 2015) and the data nature is hierarchal.

In addition to this use single level count regression model by aggregated a variables at a lower level are moved to a higher level in this case different data values from many sub-units are combined into fewer values. As the result much information is lost, and the statistical analysis loses power. On the other hand, disaggregated a variables at higher level moved to a lower level also in this case few data values from a small number of super-units are 'blown up' into many more values for much larger number of sub-units.

Ordinary statistical tests treat all these disaggregated data values as independent information from the much larger sample of sub-units. Using large number of disaggregated case for the sample size leads of significant test that reject null-hypothesis far more often than alpha level, it come up with many ‘significant’ results that are totally spurious. And also to analyzing the data and giving conclusions are mismatched at each levels. It leads to Statistical and Conceptual problems(Hox et al., 2017).

Also multilevel count regression model approach is simultaneous to analyze the effects of individual and region level of predictors and also to analyze variation on fertility status of married women between regions of Ethiopia (Goldstein, 2011).

Based on this reason multilevel count regression model approach is better relatively to single level count regression analysis, to identify the covariates related to fertility status of married women in Ethiopia.

By our findings until now, there is no any of studies on fertility status of married women’s using multilevel count regression approach and multilevel negative binomial model. But multilevel negative binomial model better than Poisson model due to excess variability, a condition called over-dispersion and use quasi-Poisson if the reverse is true.

Therefore, this study tried to identify socio-economic, demographic and geographic factors, which could contribute for the level of fertility status in Ethiopia. In this regard, the research questions of the interests were:

- ✓ What are the factors that affect fertility status of women in Ethiopia consecutively?
- ✓ Is there a significant variation in fertility status across regions in Ethiopia?
- ✓ Which count regression model is better to analyze fertility status in Ethiopia?
- ✓ Is multilevel count model is better than single level count model?

1.3. Objective of the study

1.3.1. General objective

The general objective of this study was to identifying the determinant factors of fertility status among married women’s in Ethiopia using multilevel count regression model.

1.3.2. Specific objective

- ✓ To examine the effects of various risk factors on the number of children ever born per women in Ethiopia.
- ✓ To identify factors for that may explain the variation of fertility rate among women between regions in Ethiopia.
- ✓ To select the best model among count regression model to fit fertility rate of women in Ethiopia.
- ✓ To compare single level and multilevel model to identifying determinants of fertility rate in Ethiopia.

1.4. Significance of the Study

The findings of the study will help

- ✓ The study will expect to add a body of knowledge on fertility rate and to identify its related factors.
- ✓ In addition to this the end user governmental and non-governmental organizations could take intervention measures and set appropriate plans to tackle fertility problems like to balance the number of children ever born per married women and resource between the regions, also to be useful for policy making, monitoring and evaluation different activities.
- ✓ To offer flawless information for a researchers; to set appropriate model for analysis fertility status of women and also how to use over dispersed and under multilevel count regression models.

1.5. Limitations of the Study

In this study there are some challenges that we faced. The study used data from national surveys that have inherent gaps such as absence of data on children for women age is greater than 49 attempts were made to address them arising from the fact that only surviving women aged 15-49 years were interviewed. Some variables are not included because of large number of missing values like related to HIV, smokers, age at sterilization … variables and also the seven different EDHS file data the sample size is different, due to this reason to get predictors to much difficult. In addition to this, the interaction term is not considered under this study due to convergence issue.

CHAPTER TWO

2. Literature review

This chapter presents a review of the literature on determinants of fertility status of married women's, including studies conducted in the various countries mostly focus on in Ethiopia.

2.1. Trends on fertility status of married women

Fertility means the ability to reproduce; men who are fertile and able to be a father of children and fertile women are able to get pregnant and carry their baby to full term, with a live birth nine months and 5days after conception, this all happens naturally as a result of sexual intercourse(Sherazi and Bukhari, 2019) (Morgan et al., 2016)

In sub-Saharan Africa, twenty nine countries including Ethiopia are characterized by high fertility, among this Ethiopia is the second populated country next to Nigeria(Forum, 2012)

According to the most recent Ethiopian population and housing census 2007 report, the population has increased steadily over the last three decades, from 42.6 million in 1984 to 53.5 million in 1994 and 73.8 million in 2007. There were slight declines in the population growth rates over these periods, from 3.1 percent per annum in 1984 to 2.9 percent in 1994 and 2.6 percent in 2007(Agency, 2007). But according to current World Population Review, in 2018 Ethiopia has an estimated population is approximately 107.53 million, up from 2015's estimate of 98.9 million, which ranks 14th in the world, this estimate of how many people live in Ethiopia is based on the most recent United Nations projections Ethiopian population grew annually increased fourteen-fold from 0.2 percent to 2.8 percent between 1900 and the first Population and Housing Census of 1984. The starting rate of 0.2 percent per year at the dawn of the 20th century when the country's population was estimated to be just under 12 million implied a doubling time of 140 years while the ending rate of 2.8 percent translated into a doubling time of only 25.7 years. This represents a massive increase in population on par with very high rates of growth recorded elsewhere, notably in the rest of Sub-Saharan Africa, and some Middle-Eastern countries. The Ethiopian population doubled (from the 1900 size) in sixty years (rather than 140 years) to reach 23.5 million in 1965, and doubled again in just 25 years to reach 48 million in 1990(Adugna, 2014)

The phenomenal growth was fueled by the combined impacts of increased fertility and declining mortality. The total fertility reached a high of 7.7 births in the early 1990s before coming down to 5.5 by the year 2000(Adugna, 2014).

According to the 2016 Ethiopian Demographic and Health Survey, the total fertility rate at national level was 4.6 children per women since, the total fertility rate has decreased from 5.5 children in 2000 to 5.4 children in 2005 and finally decreased 4.8in 2011 to 4.6 in 2016 children per woman. Despite slight decrement, still a single woman is having more than 4 children throughout her life(EDHS, 2016c).

Although a slight decreasing trend has shown from year to year, it is still high as compared to developed nations. Various reasons have mentioned for the reasons that kept the fertility rates still high in Ethiopia. Poverty, low level of education, economic status and less autonomy of women and traditional barriers have usually mentioned as a reason for this persistent and high fertility rate in Ethiopia(Nations, 1993)

As we see the trend of fertility decline year to year but the population growth increased. The challenge of reason's becomes increased population momentum powered by a rising population growth which continued well into the early 1990s, then leveled off at high level of 2.7 percent; Lowered mortality, and High rural fertility, which seems to have stalled at a TFR of 6.0, despite much increased contraceptive prevalence rate (CPR)(Adugna, 2014).

Estimates of the future population size and growth rate of Ethiopia with an assumption of moderate fertility decline (medium variant) indicate a substantial increase in the coming decades. Though the growth rate declines gradually, the population is expected to increase from the current 75 million to 83 million in 2010, 94.5 million in 2015 and will reach 129.1 million by 2030(EconomicDevelopment, 2006).

2.2. Empirical literature on fertility status

A research conducted by (Priya and Kshatriya, 2016) University of Delhi, on factors affecting fertility rate age group from 15-49 years among Dhangars of Madhubani District, Bihar using *One-way ANOVA* linear regression analysis revealed that age of woman, maternal age at first conception, income level, ideal number of children desired, ideal number of son desired and experience of child death were the most significant variables that explained the variance in fertility. Women who considered a higher number of children as ideal, who had a desire for son, and those who had a child death experience were more likely to have a higher number of mean live births than their counterparts. On the other hand, those who married and had their first conception at a later age, were literate, those who has a household income of more than 10000 per person and who breastfeed their children for more than 2 years had a lower number of mean live births as compared to their counterparts.

Similarly, using Multiple linear regression modeling approach a study conducted by (Berlie and Alamerew, 2018) on Determinants of Fertility Rate among Reproductive Age Women (15-49) in Gonji-Kollela District of the Amhara National Regional State, Ethiopia the result shows that sex preference, age at first birth, low educational levels of mothers and age at first sexual intercourse were the determinant factors for a high number of children ever born, Also the result was showed that early marriage, low level of formal and informal education, parent's motive to have a large number of children and inaccessible in the use of contraceptive methods were the major factors for high fertility rate in the study area.

Although, a research conducted by (Sharma, 2015) applying multiple regression analysis on the Determinants of Fertility among Women of Reproductive Age in Nepal using 2006 NDHS data, the result shows that age of respondents at first birth and educational attainment have strong negative impact on fertility; but son who have died, daughter who have died, parity at sterilization and age at sterilization have strong positive impact on fertility whereas regions, type of place of residence, age at marriage and destination India have weak positive impact on fertility.

A research conducted by(Awad and Yussof, 2017) investigates long and short term determinants of fertility rates in Malaysia based on basic macroeconomic variables for the period 1980-2014 using Auto Regressive Distributed Lag (ARDL) method. The study reveals that over a long term period, all the selected variables (GDP, infant mortality rate, females' education and employment) have had significant and negative impact on total fertility rates. Whilst during the short term period, only the infant mortality rate has had a positive impact.

A study done in Enderta District, Tigray Region, Northern Ethiopia by (Atsbaha et al., 2016) on determinants of high fertility among ever married women showed that Age at first marriage, under five child mortality, educational status of the women, current age of the women and age at first birth were found to be statistically significant of fertility status in study area using binary logistic regression model.

Similar study by applied binary logistic regression model conducted by (Abebe et al., 2018) in Arba Minch University, on determinants of high fertility among married women in Angacha district, kambeta tembero zone, southern Ethiopia, by applying case control study design the result shows that, Educational status of women, desire to have more children before marriage, age at first marriage, history of under-five mortality and not ever use of contraceptive methods were identified as determinants of high fertility. It indicates that educational status of women, age at first marriage, desire to have children before marriage, not ever use of contraceptive methods and experiencing under-five mortality were identified as determinants of high fertility.

Similarly the study by applied binary logistic regression, conducted by (Dana, 2018) on identifying demographic, socioeconomic, and cultural factors that affect fertility status among women of child bearing age from 15-49 in Wolaita-Sodo University using EDHS 2011 data; the result shows that Region, women educational level, wealth index, husband's/partner's educational level, number of visit, marital status, age at first cohabitation and age in 5-years group were found to have significant effect on total number of child ever born at 5% level of significance.

A research conducted by (Dwivedi et al., 2016) on factors affecting children ever born of women age from 15-49 in Botswana application of Poisson regression model using the 4th 2007 Botswana Family Health Survey (2007 BFHS IV) data, the result indicates that mothers employment, watch television, age of mother, place of residence, mothers education level and contraceptive use are have significant effect on total number of child ever born.

Similarly a study by applied count regression models, done by (Pandey and Kaur, 2015) on Fertility status of women (never married and ever married) between the ages of 15 and 49 in Delhi, India University using Cross-sectional data of Indian third round National Family Health Survey 2005–2006, trends and patterns of preference for birth counts suggest that religion, caste, wealth, female education, and occupation are the dominant factors shaping the observed birth process.

Similarly, by applied count regression model research conducted by (Muluneh, 2016) on Fertility Status of Married Women and its Determinants in Ethiopia using 2014 Ethiopia Mini Demographic and Health Survey data with generalized linear model (GLM),the result show that high fertility was independently associated with residing in urban areas, increased household economic status, younger age at first birth and not using contraceptives. Current age and media exposure, household head gender and media exposure, household head gender and regional state, mother's education and, regional state and media exposure and regional state were found to jointly affect fertility level.

Although, the research conducted by(Cesar Augusto Oviedo Tejada1, 2017) the socio demographic, behavioral, reproductive, and health factors associated with fertility in Brazil on to analyze the socio demographic, behavioral, and reproductive factors associated with fertility rates among Brazilian women aged between 15-49 years, by applying poison regression model to analyze data from the 2006 PNDS (Pesquisa National Demographic survey) the results show that the following characteristics are positively associated with an increase in the number of children born: age of mother, region, race, employment, year of schooling, wealth index, marital status, first sexual intercourse and age at the first birth. Thus, it is important to implement efficient family planning policies targeting these subgroups in order to improve life conditions, reduce inequalities and avoid the adverse outcomes of high fertility.

A study in Ethiopia conducted by (Ayele, 2015) using 2011 Ethiopian demographic and health survey data, by applying liner mixed model revealed that family size, age of respondents at 1st sex, education status, current age of respondent, place of residence, religion, region are associated with fertility status.

2.3. Literature review on multilevel Count Regression Models

A common model for count data is the Poisson model by assuming that the distribution has mean and variance equally(Chatfield et al., 2010). But this does not hold true in real data; a common problems to deal this model are overcoming of the sample variance is considerably larger than the mean called over-dispersion or a rare case sample variance is smaller than the mean called under-dispersion(Chatfield et al., 2010). If the data may be under dispersion to use quasi-Poisson regression provides valid inference, guarding against drawing of incorrect conclusions(Ver Hoef and Boveng, 2007, Schall, 1991).

The interpretation of the parameter estimates in the quasi-likelihood approach remains as ordinary Poisson regression. The mostly common problems to deal this model is the overcoming of over-dispersion(Chatfield et al., 2010, Iddi and Molenberghs, 2012). An over-dispersed model which assumes equal dispersion can result in misleading inferences and conclusions, as over-dispersion can lead to the under-estimation of parameter standard errors and falsely increase the significance of beta parameters(Faddy and Smith, 2011, Hilbe, 2011). As a result over-dispersed count data are common in many areas which in turn, have led to the development of statistical methodology for modeling over dispersed (Sellers and Shmueli, 2013).The negative binomial distribution looks like the Poisson distribution, but with a longer, fatter tail to the extent that the variance exceeds the mean. Depending on the degree of over dispersion, the negative binomial model can capture over dispersion than that of Poisson model(Hilbe, 2011). Multilevel models were developed to analyze nested data (Goldstein, 1995, Hox et al., 2017). In count data analysis, there are many ways in which the data may be nested: when the data can be naturally grouped. Multilevel modeling is described for analysis of correlated grouped count data. Standard models are not suitable for nested data because the independence assumption is not generally true. In order to take into account the dependence underlying the observations, random coefficients are included in the specification of the model, which is the main difference between multilevel models and the usual ones.

The multilevel count regression models assume that there is a hierarchical data set, with one single outcome or response variable measured at the lowest level, and explanatory variables at all existing levels. Conceptually, it is useful to view the multilevel regression model as a hierarchical system of regression equations (Hox et al., 2017)

CHAPTER THERE

3. Data and methodology

3.1. Source of Data

Data source of this study was use secondary data obtained 2016 Ethiopian Demographic and Health Survey (EDHS, 2016) data is the national, population-based, cross-sectional survey. Is the fourth survey implemented by the Central Statistical Agency (CSA). It was conducted from January 18, 2016 to June 27, 2016, included a total of 18,008 households selected for the sample, of which 17,067 were occupied. Of the occupied households, 16,650 were successfully interviewed, yielding a response rate of 98% (EDHS, 2016a).

The EDHS 2016 sample was a complex sampling design (i.e. combined stratified and cluster in two stages, with unequal probabilities of selection that result in weighted sample to separate the sample components) and was designed to provide reliable representative estimates of the national, and regional level administratively.

The survey collected a total of 15,683 successfully interviewed women aged from 15-49 years out of this 9,602 women were considered in this study.

3.2. Variable include in the study

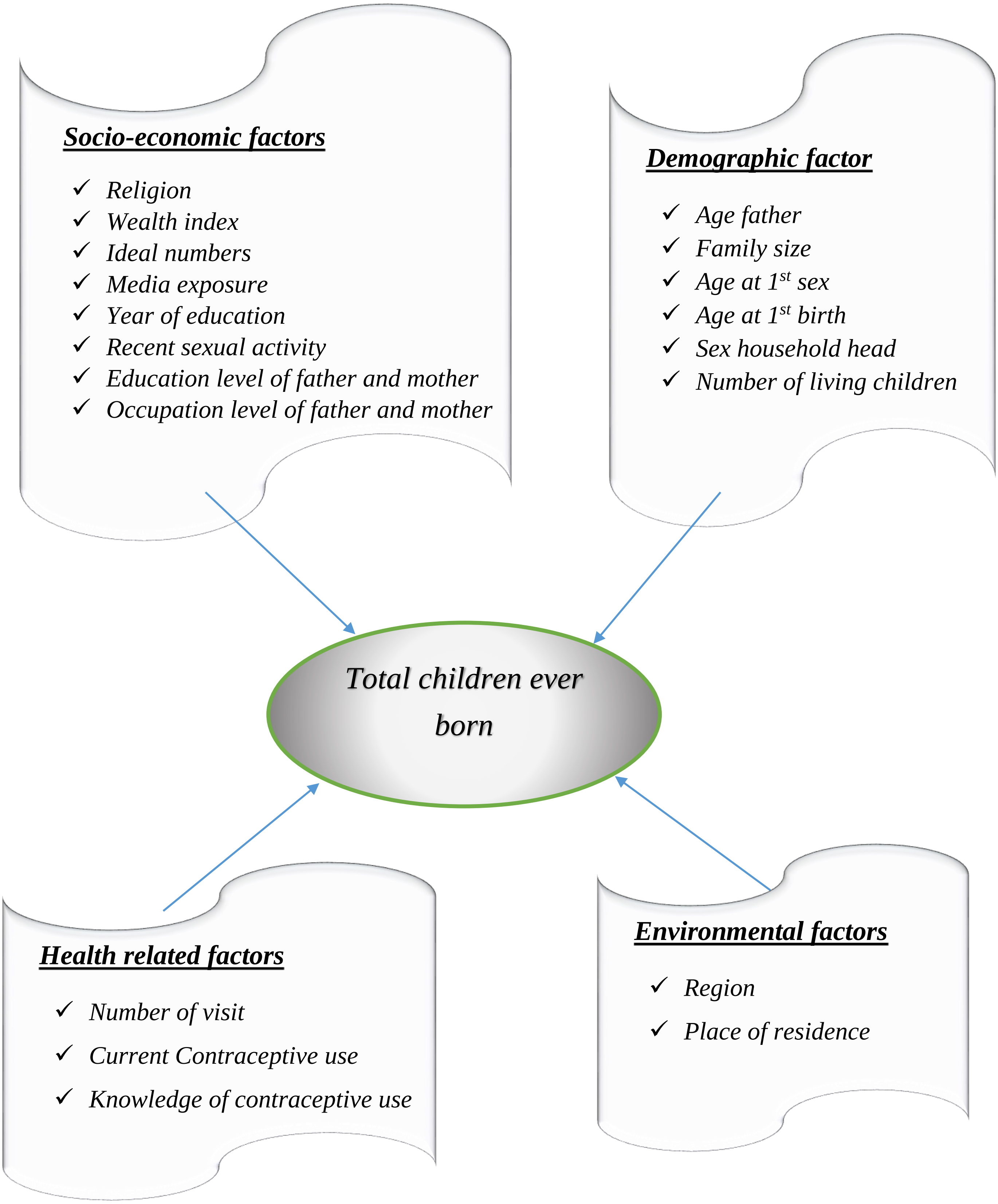
3.2.1. Response variable

Response variable denote (Y_i) is total number of ever born children per married women in Ethiopia over her lifetime fertile up to the survey date.

Where, Y_i=0, 1, 2, 3… *i.e.* i refers to the individual mother.

3.2.2. Predictor/explanatory variables

Based on literatures, several variables that are associated with fertility status were include as explanatory variable, thus are follow as below:


Socio-economic factors
Religion
Wealth index
Ideal numbers
Media exposure
Year of education
Recent sexual activity
Education level of father and mother
Occupation level of father and mother
Demographic factor
Age father
Family size
Age at 1st sex
Age at 1st birth
Sex household head
Number of living children
Total children ever born
Health related factors
Number of visit
Current Contraceptive use
Knowledge of contraceptive use
Environmental factors
Region
Place of residence

3.3. Methods of data analysis

Count regression is a popular modeling approach when the dependent/outcome variable is integer or count data(Chatfield et al., 2010). These models can be employed to examine occurrence and frequency of occurrence. The most popular distribution to model count data is the Poisson distribution, which is based on the property that the mean and variance of the dependent variable are assumed to be equal(Chatfield et al., 2010). However, it's not always happened, mostly variance exceeds then mean. It indicates to over dispersion(Carlin et al., 2014, Cameron and Trivedi, 2013). Over dispersion can be modeled 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. If the reverse is happened to apply quasi-Poisson.

In this study single and multilevel count regressions are employed to identify determinant factors of fertility. Firstly, we analyzed the data using single level count regressions by assuming the occurrence of fertility status is independently. And finally we assessed the determinants and regional difference on fertility status (total number of ever born children) using multilevel count regression model.

3.3.1. Single level count regression model

3.3.1.1. Poisson regression:

Let Y_1, .$Y_{2,}$ . . . . Y_n be independent random variables with Y_i denoting the number of events observed from exposure n_i for the i^{th} covariate pattern (Goldstein, 2011).
The probability mass function of Poisson distribution is given as:

$$P\left(Y = \frac{yi}{\mu i}\right) = \frac{\exp(-\mu i)\mu i^{yi}}{yi!} \quad where\ \mu > 0\ , yi = 0,1,2 \ldots n, i = 0,1,2, \ldots \ \ldots\ldots\ldots\ldots\ldots\ (3.1)$$

$\mathbf{\mu}i$, *is* the average number of occurrences of an events

Now, $\mathbf{y}i$ is total number of ever born children in the i^{th} mother at a given time with parameter $\mathbf{\mu}i$ then the expected/mean and variance of $\mathbf{y}i$ can be written as:

$$Var(Yi) = \mu = E(\text{yi}) \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (3.2)$$

If we write Poisson probability mass function pmf in the form of exponential family distribution

$$\left.\begin{aligned} f(y,\theta,\phi) &= \exp(a(y)b(\theta)+c(\theta)+d(\mathrm{y}), \\ f(y,\theta,\phi) &= \exp[\{\tfrac{y\,\theta-b(\theta)}{a(\phi)}\}+c(y,\phi)] \end{aligned}\right\} \text{ or} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.3)}$$

(y) is a general form of Exponential Family Distribution then the Poisson distribution is written as:

$$\left.\begin{aligned} f(y,\mu,\phi) &= \exp(y\log\mu-\mu-\log y!) \\ f(y,\mu,\phi) &= \exp[\{\tfrac{y\log\mu+\mu}{a(\phi)}\}-\log y!] \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.4)}$$

The expected and variance of the distribution is: $E(Y) = b'(\theta)$

$V(Y) = a(\phi) * b''(\theta)$

Using the 1st form:

Then, $a(y) = y$, $b(\theta) = \log\mu$, $c(\theta) = -\mu$ and d(y)$= \log y!$

Using the 2nd form:

$\theta = \log\mu \Rightarrow \mu = e^{\theta}$, and $b(\theta) = \mu = e^{\theta}$ and $a(\phi) = 1$

The mean and variance of Poisson distribution is given by:

$b'(\theta) = e^{\theta} = \mu = E(Y)$ and $V(Y) = a(\phi) * b''(\theta) = \mu$

The estimation is undertaken by using maximum likelihood method. The likelihood function of the Poisson model based on a sample of n independent observations is given by

$$l(\mu,y) = \prod_{i=1}^{n} \frac{\exp(-\mu)\mu^{y}}{y!} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3.5)}$$

The log-likelihood function for Poisson distribution is.

$$L(\mu,y) = \log\big(l(\mu,y)\big) = \sum_{i=1}^{n}(yi\log\mu i-\mu i-\log yi) \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(3.6)}$$

The likelihood equation for estimating the parameter is obtained by taking the partial derivations of the log-likelihood function and setting them equal to zero.

Now we can link the Poisson distribution of natural parameter with the predictors Let as X is a $n \times (p+1)$ matrix of explanatory variables and $\beta\ be\ (p+1) \times 1$dimensional column vector of unknown parameters to be estimated. Now we can link the Poisson distribution of natural parameter with the predictors as:

$$\log(\mu_i) = X'\beta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots \ldots \beta_k X_k \quad \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (3.7)$$

The model is known as Poisson regression/log-linear model implies that the logarithm of the response variable is linear relationship with predictor's means that the change in the log of the response variables is linear with change in the explanatory variable. Equivalently; easy to interpret the log-linear model is written as: $\mu_i = \exp(X'\beta)$

In testing over dispersion, the hypothesis is given by:

$H_{O:}$ α=1 VS $H_{1:}$ α>1

There are two basic criteria commonly used to check the presence of over dispersion:

1. Deviance test

$$D(y, \mu^{\wedge}) = 2 * \sum_{i=1}^{n} (yi \log(\frac{y_i}{\mu^{\wedge}}) - (y_i - \mu^{\wedge})) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (3.8)$$

Where, y is the number of events, n is the number of observations and $\mu^{\wedge}$is the fitted Poisson mean.

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

$$x^2 = \sum_{I=1}^{n} \left(\frac{(y_i - \mu^{\wedge})}{\mu^{\wedge}} \right) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (3.9)$$

If the model fits the data, both deviance and Pearson Chi-square statistics divided by the degrees of freedom are approximately equal to one. Values greater than one indicate the data is an over dispersed, while values smaller than one indicate an under-dispersion. Another test for over dispersion is LRT. If P-value of LRT α < (level of significance), then there is over-dispersion and the Negative Binomial model is preferable.

3.3.1.2. Negative binomial regression model

The negative binomial model is an extension of the Poisson model to overcome possible over dispersion in the data caused by heterogeneity or an excess number of zeros(Chatfield et al., 2010). If a Poisson regression model doesn't fit the data and it appears that the variance of Y_i's increasing faster than the mean, then a simple scale-factor adjustment is not appropriate. One way to handle this situation is to fit a parametric model that is more dispersed than the Poisson. A natural choice is the negative binomial(Chatfield et al., 2010) 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(Cameron and Trivedi, 2013, Harris et al., 2014).

The NB regression model is given by

$$p(y_i, \mu_i \alpha) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i \Gamma\left(\frac{1}{\alpha}\right)} \left(\left(1 + \alpha\mu_i\right)^{\frac{-1}{\alpha}}\right) \left(\left(1 + \frac{1}{\alpha\mu_i}\right)^{-y_i}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.10)$$

Where $y_i \geq 0$ and $\alpha > 0$

Where α is the over dispersion parameter and $\Gamma(.)$ is the gamma function when α=0 the Negative Binomial distribution is the same as Poisson distribution. The mean and variance are expressed as:

$E(y_i) = \mu_i = exp(x_i'\beta)$ And

$Var(y_i) = \mu_i\,(1 + \alpha\mu_i)$

NB GLM, the mean response for the number of ever born children's per mother is assumed to have a log -linear relationship with the covariates and is structured as:

$$\log(\mu_i) = x_i'\,\beta$$

Where, x_i= selected determinant factor of number of ever born children's and β represents regression coefficients to be estimated.

The NB likelihood function is

$$l(\mu_i, \alpha, y_i) = \prod_{i=1}^{n} \left[\frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_i \Gamma\left(\frac{1}{\alpha}\right)} \left((1 + \alpha\mu_i)^{\frac{-1}{\alpha}}\right) \left(\left(1 + \frac{1}{\alpha\mu_i}\right)^{-y_i}\right)\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.11)$$

The log-likelihood function for NB regression model is

$$L(\mu_i, \alpha, y_i) = \log\big(l(\mu_i, \alpha, y_i)\big)$$

$$= \sum_{I=1}^{N} \left[-\log Y_I! + \log\left(\frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)}\right) - \frac{1}{\alpha}\log(1 + \alpha\mu_i) - Y_I \, \log\left(1 + \frac{1}{\alpha\mu_i}\right)\right] \ldots\ldots.. (3.12)$$

The likelihood equations for estimating μi and α are obtained by taking the partial derivations of the log-likelihood function and setting them equal to zero.

Furthermore, negative binomial model can solve an over-dispersion problem, it may not be handle when there are excess zeros in the data; in such cases zero -inflated models are preferable.

3.3.2. Multilevel regression model:

Before going to multilevel model first we talk about multilevel data, is any kinds of data including observational data collected in the human and biological sciences have *clustered or nested/hierarchical* structure. Hierarchies are one way of representing the dependent or correlated nature of the relationship between individuals and their groups(Maas and Hox, 2002).

Multilevel regression model is simply an extension of the classical multiple regression model, it introduce additional parameters, (i.e. additional number of residual variance across a level, random coefficient (intercept and slope coefficient's are assumed to vary across the level) and interaction effect between variables) on the model with nested nature structure of the data. They all assume that there is a hierarchical data set, with one single outcome/response variable that is measured at the lowest level, and explanatory variables may existing at all levels(Maas and Hox, 2002).

Multilevel analysis is a methodology for the analysis of data with complex patterns of variability, with a focus on nested sources of variability. The best way to the analysis of multilevel data is an approach that represents within group as well as between group relations within a single analysis, where group‟ refers to the units at the higher levels of the nesting hierarchy(Maas and Hox, 2002).

3.3.2.1. Multilevel Poisson regression model

The Multilevel count model have been described by(Bock and Mislevy, 1989). In generalized multilevel models, the multilevel structures appear in the linear regression equation of the generalized linear model. Simply Multilevel count model is an extension of the generalized linear model that includes nested /hierarchy data structure and random coefficients.

The appropriate approach to analyzing fertility status of married women's is Multilevel count regression model, since the structure of data in the population is hierarchical, and a sample from such a population can be viewed as a multistage sample with a single integer response/outcome variable. The 2016 EDHS data set used for this study is based on a multistage stratified cluster sampling. Here the units at lower level are individuals (married women aged from 15 *to* 49) who are nested within units at higher level (regions).

In multilevel modeling, Model building strategies can be either top-down or bottom-up. But in multilevel modeling, bottom-up strategy better than top-down strategy in order to start simple model or go to step by step to prevent complicated model.

Based on this we start the simplest model one which is called *intercept-only model.*

1. *Intercept-only* model *is a model with no explanatory variables* contains only a response variable and an intercept (Hox et al., 2017), which is written as:

The link function of poison distribution is logarithm, then $\eta_{ij} = \log(\mu_{ij}) = \ln(\mu_{ij})$

$\eta_{ij} = \beta_{0j}$, $\beta_{0j} = \gamma_{00} + u_{0j}$ then intercept only model becomes:

$$\eta_{ij} = \gamma_{00} + u_{0j} \quad \text{………………………………………………………} (3.13)$$

<u>Where</u>

β_{0j} Is refers that intercept of the dependent variable in group j (Level 2) which is region

γ_{00} is refers to the overall intercept. Which is the grand mean of the scores on the dependent variable across all the groups when all the predictors are equal to 0.

u_{0j} Is a random error component of variation at group level, measures regional variation of fertility status.

2. Analyze a model with all lower-level explanatory variables are fixed. This means that the corresponding variance components of the slopes are fixed at zero.

If there are p explanatory (i.e. $X1, X2, \dots\dots Xp$) variable at lower-level which introduce in model. Then the model is written as:

$$\eta_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + u_{0j} \quad \text{……………………………………………} (3.14)$$

Where, Xij are a variable with p explanatory variables at the individual-level and γ_{p0} refers to the slope of regression coefficient between the dependent variable and the Level-1 predictors.

This step to be use

- ✓ To assess the contribution of each first-level explanatory variable on the model.
- ✓ Identify significance variables of each predictor and to observe changes of second-level variance terms.
- ✓ And also we can compare the first model (intercept-only model) and the second model (the model with Lower-level explanatory are introduce).

3. Random slope model is a model Lower-level explanatory are random (not fixed) on the groups implies that slopes for each regression coefficients are vary across regions. Then the model is written as:

$$\eta_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + (u_{pj})X_{pij} + u_{0j} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.15)$$

Here we assess whether any of the slopes any explanatory variables has a significant variance component between the groups (i.e. the slope of lower-level predictor's are vary across regions).

Another meaning the relationship between an explanatory variable and the response is not the same across all regions. If we fit a model based on the same predictors on the response variable for all regions separately, we may obtain different slopes for each region.

4. Random intercept model is a model in which intercepts are allowed to vary, therefore, the scores on the dependent variable for each individual observation are predicted by the intercept that varies across regions. That means the groups differ with respect to the average value of the response variable, but the relation between explanatory and response variables can't differ between groups. The random intercept model expresses the natural log of λ_{ij} as a sum of a linear function of the explanatory variables. That is

$$\eta_{ij} = \gamma_{0j} + \gamma_{p0}X_{pij} + u_{0j} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.16)$$

5. Random coefficient model is a model both the intercepts and slopes are vary across a region, the model becomes:

$$\eta_{ij} = \gamma_{0j} + \gamma_{p0}X_{pij} + (u_{pj})X_{pij} + u_{0j} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.17)$$

Mostly in count response variance is excess than the mean we call it over dispersed, the obvious solution to overcome this problem is using fixed offset variable may be better to center them about their mean in order to avoid numerical instabilities. Then the full model equation for the two-level Poisson regression with i^{th} individual mothers are nested within the j^{th}region the model becomes:

$$\eta_{ij} = \underbrace{\log(w_{ij}) + \gamma_{00} + \gamma_{p0}X_{pij}}_{\text{Fixed effect of the model}} + \underbrace{(u_{pj})X_{pij} + u_{0j}}_{\text{random effect of the model}} \ldots\ldots\ldots\ldots\ldots\ldots (3.18)$$

Where, $\log(w_{ij})$ is an offset variable (Desire number of children) with fixed $intercept$ $term\ \gamma_{00}$.

Or the model with random intercept term is written as:

$$\eta_{ij} = \underbrace{\log(w_{ij}) + \gamma_{p0}X_{pij}}_{\text{Fixed effect of the model}} + \underbrace{\gamma_{0j} + (u_{pj})X_{pij} + u_{0j}}_{\text{random effect of the model}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.19)$$

The assumption of Poisson distribution for the observed count η_{ij} which are assumed conditionally independent with $E(\eta_{ij}) = \mu_{ij}$, $Var(\eta_{ij} \mid \mu_{ij})$=$\mu_{ij}$

If to estimate additional parameter which is offset variable the value of mean and variance still unequal implies there is over-dispersion then multilevel Negative Binomial model is preferred.

3.3.2.2. Multilevel negative binomial (NB) regression model

NB regression model is a popular approach to analysis over-desperation data .Often, because of the hierarchical study design or the data collection procedure, over-desperation and lack of independence may occur simultaneously, which render the standard NB model inadequate. To account for the over-desperation and the inherent correlation of observations, a class of multilevel NB regression model with random effects is presented.

The multilevel NB model

$$\log(\mu_{ij}) = \eta_{ij} + e_{ij} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.20)$$

Where, $\eta_{ij} = \gamma_{0j} + \gamma_{1j}x_{1ij} + \gamma_{2j}x_{2ij} + \cdots + \gamma_{kj}x_{kij} \Rightarrow \gamma_{0j} + \gamma_{pj}x_{pij}$

$cov\ (\eta_{ij}, e_{ij}) = 0$

$Exp(e_{ij})$ follows gamma probability distribution,$\Gamma(v)$, with mean 1 and variance $\alpha = v^{-1}$. Integrating with respect to e_{ij} (Cameron and Trivedi, 1986) the resulting probability distribution.

$$P(Y_{ij} = y_{ij}) = \frac{exp\left(-exp(\eta_{ij} + e_{ij})\right) exp(\eta_{ij} + e_{ij})^{y_{ij}}}{y_{ij}!} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.21)$$

One version of the multilevel negative binomial regression model is obtained;

$$P(Y_{ij}=y_{ij})=\frac{\Gamma(y_{ij}+v)v^{v}{\mu_{ij}}^{y_{ij}}}{y_{ij}!\,\Gamma(v)(v+\mu_{ij})^{v+y_{ij}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.22)$$

With mean $\mu_{ij}=\exp(\eta_{ij})$ and variance, $Var(y_{ij})=\mu_{ij}+\alpha{\mu_{ij}}^{2}$ where, α is the dispersion parameter.

3.4. Parameter estimation

3.4.1. Maximum likelihood

Maximum Likelihood (ML) is the most commonly used estimation method in multilevel modeling. An advantage of the Maximum Likelihood estimation method is that it is robust, Produce estimate that are asymptotically efficient and consistent; with large sample size, ML estimate are usually robust against mild violations of the assumption, such as having non- normal errors.

Maximum Likelihood estimation proceeds by maximizing a function called the likelihood function. The two Likelihood function are used in multilevel regression modeling. One is called full Maximum Likelihood (FML); in this method both regression coefficients and variance components are included in the likelihood function.

The other method is called Restricted Maximum Likelihood (RML); here only variance components are included in the likelihood function and the regression coefficients are estimated in a second estimation step. Both methods produce parameter estimates with associated standard error and overall model deviance which is function of likelihood.

Approximate MLEs of GLM we rely on Newton-Raphson algorithm. The log- likelihood functions of β.

The EFD form GLM is given by:

$$f(y,\mu,\phi)=\exp\left(\frac{y_i\theta_i-b(\theta_i)}{a(\phi)}-c(y_i,\phi)\right)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(3.23)$$

Let us consider a random sample of size n, yi where, $i=1,2,\ldots.n$

Then the likelihood function can be written as:

$$L(\theta,Y,\phi)=\prod_{i=1}^{n}\left(f(y,\mu,\phi)\right)=\prod_{i=1}^{n}\exp\left(\frac{y_i\theta_i-b(\theta_i)}{a(\phi)}-c(y_i,\phi)\right)\ \ldots\ldots\ldots\ldots\ldots (3.24)$$

Moreover, the log likelihood function is:

$$L(\theta, Y, \phi) = \sum_{i=1}^{n} \left(\frac{\mathrm{y_i \theta_i - b(\theta_i)}}{\mathrm{a}(\phi)} - c(y_i, \phi) \right)$$

$$= \sum_{i=1}^{n} l_i(\theta_i, \phi) = \sum_{i=1}^{n} l_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.25)$$

The MLE of β are the solve using derivate with respect to β equal to zero.

3.4.2. Generalized least squares

Generalized least squares (GLS) estimate can be obtaind from ML procedure by restricting the number of iteration to one. In large sample size GLS and ML are equivalent.

3.4.3. Generalized Estimating equation

Generalized Estimating equation (GEE) method estimate the variance and covariance in the random part of multilevel model directly from resduals, which makes them faster than to compute than full ML estimate. After variance components obtaind, GLS is used to estimate the fixed regresssion cofficients. And GEE used to estimate population average model,where comparing the group level units.

GEE estimate are different from MLE estimate when nonliner model is estimated. When heteroscedasticity is involved due to non-normality ,outliers or misspecification of the model, asymptotic standard error are small in this case apply GEE, otherwise use MLE.

Note: GEE is better as compared to MLE since a sample variances distribution is skewed.

3.5. Assessing model fit

3.5.1. Test for significance of a single variable

3.5.1.1. Wald test

Wald test is used to test statistical significance of each coefficient (β) in the model. The null hypothesis for this test may be stated as "Factor x_i does not have any value added to the prediction of fertility status given that other factors are already included in the model."

H_0: β_j=0 vs H_1: $\beta_j \neq 0$

To test such a null hypothesis, $z = \frac{\beta j}{s.e(\beta_j)}$

Z, Has an approximate standard normal distribution. Equivalently, Z^2 has approximately a chi-squared distribution with one DF.

In multilevel, it is not appropriate for variances, because it assumes normal distribution,while the sampling distributionof variances is skewed. They propose to use chi-square test of the residual. This chi-square is computed as: $X^2 = \sum \frac{(\beta^\wedge{}_j - \beta)^2}{V^\wedge{}_j}$

where $\beta^\wedge{}_j$ is the regression coefficient computed separately in group j, β is overall estimate and $V^\wedge{}_j$ is estimated sampling variance in group j and also the number of degree of freedom is given by *df= j-p-1*, where j is a number of second level units,and p is a total number of explanatory variables in the model.

3.5.2. Test for significance of overall model

3.5.2.1. Likelihood ratio test

For nested models, the LRT is a test of a null hypothesis H_0 against an alternative H_1 based on the ratio of two log-likelihood functions. The likelihood ratio test is a test of the overall model. The overall test statistic for likelihood ratio test is given as(Lawless, 1987): Likelihood ratio test statistics is given by .

$$G^2 = -2(L_0 - L_1) \sim X^2(p-1)$$

Where $L_0$the log-likelihood of the null is model and L_1 is the log-likelihood of the full model. This method is not appropriate for models which are non-nested one on the other. For this study LRT is used to compare the Poisson with the negative binomial. Since, Poisson is nested on negative binomial. we used method such as the Akakie information criteria (AIC) and Bayesian information criteria (BIC) to compare non nested model (Ayalew et al., 2014).

3.5.3. Model comparision

3.5.3.1. Comparing nested models

Nested model: the level of one factor only make sense with in the level of another factor. For nested model all term of smaller model occur in larger model it is necessary condition for using model comparison tests like likelihood ratio test (*deviance*).

From the likelihood function we can calculate a statistic called the *deviance* that indicates how well the model fits the data. The *deviance* isdefined as:

D=**-2 Log (Likelihood)** $D=-2\{log\hat{L}c - log\hat{L}f\} = -2\{log\{\hat{L}c/\hat{L}f\}$ …………...…….. (3.26)

Where $\hat{L}c$ and $\hat{L}f$ are the maximized log-likelihood of models under the null and alternative a hypothesis respectively.
Note: Models with a lower deviance fit better than models with a higher deviance.

3.5.3.2.Comparing non-nested models

Non-nested model: factor is combination of two factors that are not related.

- ✓ Bayesian information criterion (BIC) is defined as:

  BIC=d+q Log (Likelihood),

  where q is number of estimated paramter,and d is deviance

- ✓ Akaike's information criterion (AIC) is defined as:
  AIC = d+2q

In Multilevel data have different sample size at different level,in this case the AIC is more straightforward than the BIC.

Note: The model having lower AIC and BIC values is better fit than the model having higer AIC and BIC values.

3.6. Data analysis to used soft wares

The statistical analyses were performed using South Texas Art Therapy Association (STATA) version 14, Statistical Package for Social Science (SPSS) version 21, R version 3.6.0 and Microsoft-Excel.

CHAPTER FOUR

4. Result and discussion

4.1. Descriptive statistics

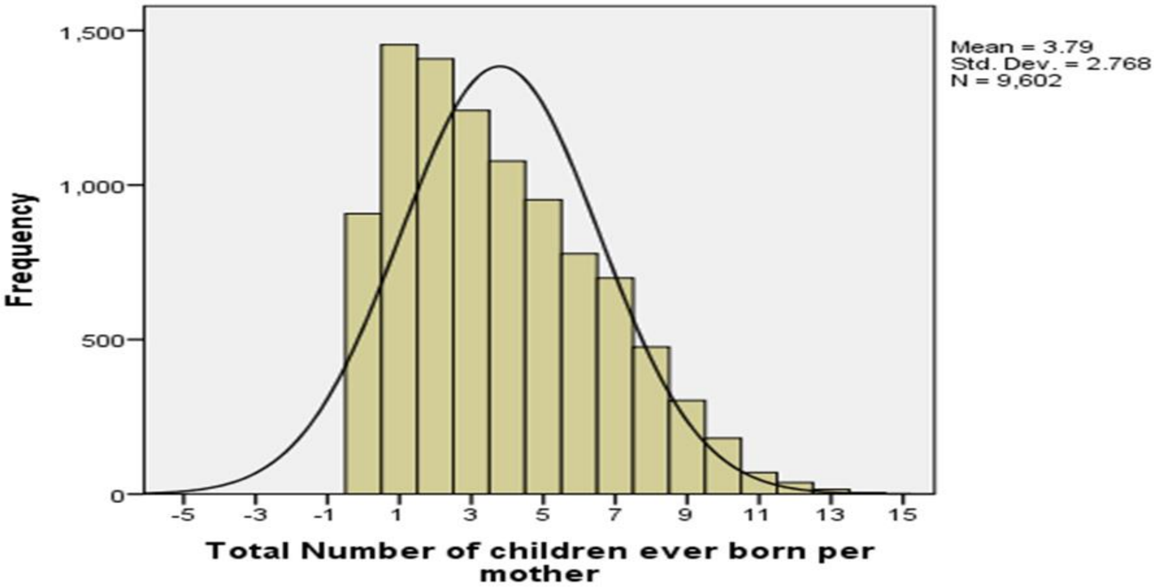
The result showed that, positively right skewed distribution. Further screening a variance of total number of children ever born calculated as (7.660) is greater than the mean (3.794) indicating over-dispersion. This is an indication that the data could be fitted better by count data models which takes into negative binomial model (Table 1, Fig. 1).

Table 4. 1. Frequency distribution of total number of children ever born per women

Total number of children ever born	Frequency	Percent	Cumulative Percent
0	907	9.45	9.45
1	1454	15.14	24.59
2	1408	14.66	39.25
3	1241	12.92	52.18
4	1077	11.22	63.39
5	952	9.91	73.31
6	778	8.10	81.41
7	699	7.28	88.69
8	476	4.96	93.65
9	303	3.16	96.80
10	181	1.89	98.69
11	70	0.73	99.42
12	37	0.39	99.80
13	15	0.16	99.96
14	4	0.04	100.00
Total	9602	100	

Mean	Variance	Minimum	Maximum	Skewness	Kurtosis
3.794	7.660	0	14	0.624	-0.277

Figure 4.1. Histogram of total ever born child per mother



Some of the socioeconomic, demographic, health and environmental related factors on the total number of children ever born per mother are summarized in Table 4.2.

On average highest number of children are born in Somali around (4.77) and lowest in Addis Ababa (1.99) as compared to 9 regions and 2 administrative city. In addition the mean number of children ever born for rural areas (4.22) is higher than or almost two-ways of urban areas (2.50).

According to highest educational level of mother's, the mean number of children ever born for uneducated level of mother (4.80) is higher than mother's with primary, secondary and above education levels (2.71, 1.73) respectively. Similarly, the mean number of children ever born for uneducated father (4.61) is greater than fathers with primary, secondary and above education level (3.66, 2.28) respectively.

According to religion of respondent's, the mean number of children ever born for believer of Muslim mother's (4.14) is higher than mother's with Orthodox, protestant and others believer's (3.83, 3.34) respectively.

Also the mean number of children ever born for male household head's has (3.89) is higher than household head for female (3.38). In addition to this the mean number of children ever born for lowest wealth index (4.39) is higher than middle and higher wealth index (4.14, 2.60) respectively.

Age of mother's at first birth the mean number of children ever born for at age lies between 16 up to 19 (4.53) is higher than age lies between 20 up to 40 and less than or equal to 15 (3.74, 3.22) respectively. Also mean number of children ever born for don't have knowledge about contraceptive use (4.53) is higher than having knowledge contraceptive use (3.74). Similarly mean number of children ever born for not using contraceptive use (4.04) is higher than using contraceptive use (3.22). In addition to this the mean number of children ever born for active recent sexual activity (3.80) is higher than not active (3.78), it's approximately the same.

According to occupation of mother's, the mean number of children ever born for not working mother's (3.85) is higher than having work mother's (3.72). Similarly the mean number of children ever born for did not work father's (4.30) is higher than did have work (3.73). Also the mean number of children ever born for not using media exposure (4.33) is higher than using media exposure (2.98).

Table 4.2. Summary statistics of predictor variables related to total children ever born

Variables	Categories	Mean	Std.Dev	Number of Observation
Region	Tigray	3.86	2.806	935
	Afar	3.77	3.023	858
	Amhara	3.73	2.683	1114
	Oromia	4.23	2.792	1286
	Somali	4.77	3.046	973
	Benishangul G.	4.11	2.773	791
	SNNPR	4.27	2.713	1198
	Gambela	3.12	2.201	686
	Harari	3.14	2.436	568
	Addis Ababa	1.99	1.495	625
	Dire Dawa	3.20	2.665	568
Residence	Urban	2.50	2.159	2369
	Rural	4.22	2.814	7233
Highest educational level of mother	No education	4.80	2.736	5625
	Primary	2.71	2.293	2621
	2^{n}ary and above	1.73	1.523	1356

Religion	Orthodox Muslim protestant and others	3.34 4.14 3.83	2.617 2.907 2.593	3416 4261 1925
Sex of household head	Female Male	3.38 3.89	2.753 2.762	1772 7830
Wealth index combined	Lowest Middle Highest	4.39 4.14 2.60	2.901 2.743 2.246	2880 4105 2617
Age of respondent at 1st birth	<=15 16-19 20-4o	3.22 4.35 3.43	3.270 2.537 2.326	2265 4127 3210
Knowledge of contraceptive use	No Yes	4.53 3.74	3.014 2.692	417 9185
Current contraceptive use	No Yes	4.04 3.22	2.893 2.354	6715 2887
Recent sexual activity	Not active Active	3.78 3.80	2.749 2.774	2342 7260
Father's education level	No education Primary Secondary and above	4.61 3.66 2.28	2.842 2.624 2.043	4454 3002 2146
Father's occupation	Did not work Did have work	4.30 3.73	3.028 2.726	1097 8505
Mother's occupation	Not working Has work	3.85 3.72	2.794 2.734	5273 4329
Media Exposure	No Yes	4.33 2.98	2.796 2.515	5794 3808

4.2 Single-level Count Regression Analysis

4.2.1 Variable Selection method

In order to select variables to be included in multivariable analysis, stepwise variable selection was used. The result recognized that: Age of mother, region, residence, religion, family size, age of mother at first birth, number of living children, occupation of father, Wealth index, year of educational level of mother, number of visits, age of mother at first sex, ideal number of children and contraceptive use are statistically significant and important variable, but the other variables are found to be non-significant and thus excluded from the analysis.

4.2.2 Goodness of Fit and Test of over dispersion

The dispersion parameter was tested to the existence of over dispersion or not on the hypothesis. H0: $\alpha=0$ (no over dispersion in the data set) vs H1: $\alpha\neq 0$(there is over dispersion in the data set).

Since, the result putted from in appendix the likelihood ratio statistics $\{-2\ [(-15749.6991) - (-15244.9945)]\} = \{(31499.3982) - (30489.989)\} = 1009$ with p-value =0.0001, we reject the null hypothesis indicted that there was over-dispersion problems and the negative binomial model more appropriate than the Poisson model.

Table 4.3. Test of over-dispersed

Criterion	Model	Value	p- value
LRT	NB	1009	0.0001

4.2.3 Model Selection Criteria

4.2.3.1. Information Criteria's and LRT values

Several model selection methods have been proposed in the literature. The most commonly used methods include information and likelihood based criteria. Described to fit the data using AIC and BIC values for each model are presented in Table 4.4.

Negative binomial regression model is more appropriate than Poisson regression model for describing a total number of ever born children per mother, since the value of Poisson regression model had the largest AIC and BIC demonstrating a poor fit to the data as compared to NB regression model.

Table 4.4. Single level count regression model selection criteria

	Criteria	
Model	**AIC**	**BIC**
Poisson	31553	31747
Negative binomial	30546	30747

4.2.3.2. Predicted value and Probability

The result showed in table 4.5 and figure 4.2 showed that predicted values of the total children ever born, Negative Binomial model were closest to the observed probabilities. Therefore the Negative Binomial regression model is more preferred model than the Poisson regression model.

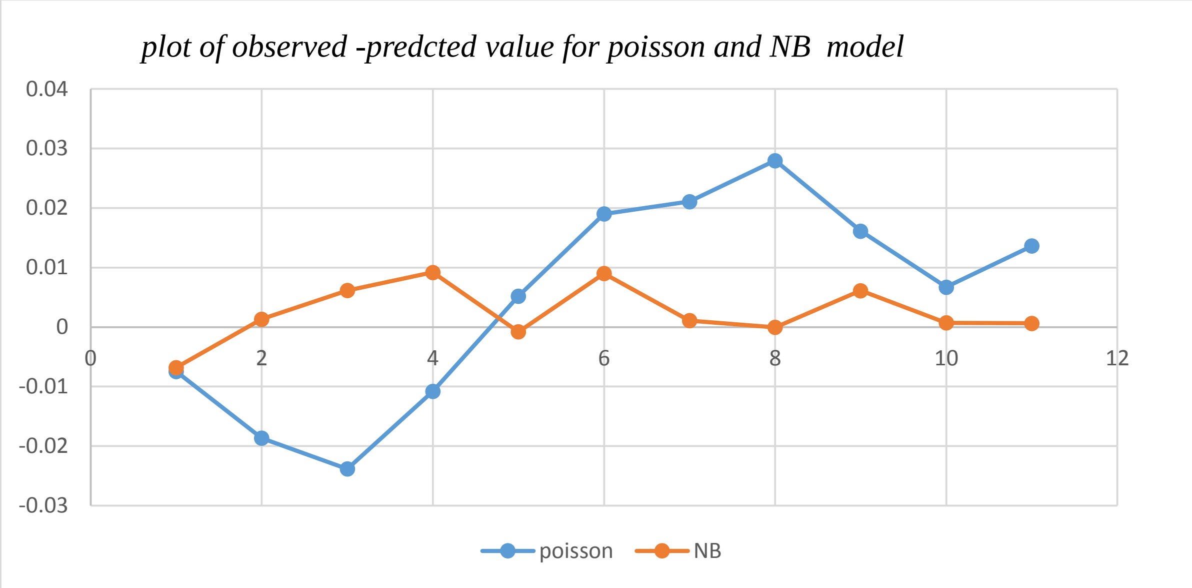
Table 4.5. Observed and predicted probability forever born children's

Total NO_ of ever born Children	Observed Probability	Predicted probability **Poisson**	**NB**
0	0.094459	0.101928	0.091245
1	0.151427	0.170102	0.150103
2	0.146636	0.170475	0.140475
3	0.129244	0.140051	0.120050
4	0.112164	0.106964	0.112963
5	0.099146	0.080141	0.090140
6	0.081025	0.059937	0.079936
7	0.072797	0.044827	0.072826
8	0.049573	0.033444	0.043444
9	0.031556	0.024832	0.030832
10^+	0.031973	0.018330	0.031330

4.2.4.3. Plots of Differences between Observed and Predicted value

Also, the result showed in figure 4.2 showed that predicted values of the total children ever born for Negative Binomial model were closest to the observed probabilities. Therefore the Negative Binomial regression model is more preferred model than the Poisson regression model.

Figure 4.2. Plots for difference between observed and predicted values



4.2.4. Parameter Estimation of Single Level Negative Binomial Model

According to table 4.6 results of NB model; wealth index, year of education and age of mother at 1st sex are negatively associated with total number of ever born children. However, family size, age of mother, age of mother at first birth, number of living child's, father's occupation, region, ideal number of children and number of visits are positively associated with a total number of ever born children.

In general the 95% confidence intervals (CI) is interpreted as includes 1 then the result is non-significant and the reverse is true, similarly the relative ratio interpreted as the mean number of children ever born at the given category times as compared to the mean number of children ever born from the reference category.

To observe the region states, is a significant factor of total children ever born per mother in Ethiopia. The mean number of total children ever born is 1.100, 1.075, 1.069 and 1.072 times higher among total ever born child in Somali, Benishangul-Gumuz, SNNP and Gambela regions as compared to total ever born child in Tigray respectively.

The mean number of total ever born child for Muslim believer's mothers has 1.063 times greater than that of orthodox believer mothers. Similarly family size increase by one unit the mean number of total ever born child is 1.014. Similarly, wealth index of the respondents has significant negative impact on the number of ever born child per mother.

Implies that the expected number of ever born child for middle and highest wealth index are lower than 0.980 and 0.937 times respectively as compared lowest wealth index. Also occupation of father has significant positive impact on the number of ever born child per mother. Means the expected number of ever born child of farther who has job attachment is higher than 1.071 times that of father who has job attachment. Similarly the reaming significant variable can be interpreted as this way.

Table 4.6. Parameter estimation of single level Negative binomial model

	Estimation	S.E	Z-value	p-value	IRR	95%CI for IRR	
						Lower	upper
Intercept	0.1203	0.0465	2.587	0.0097	1.128	1.030	1.235
Current age of mother	0.0225	0.0008	28.125	<.0001	1.022	1.020	1.024
Region (Ref=Tigray)							
Addis Ababa	-0.0116	0.0374	-0.309	0.7572	0.988	0.918	1.063
Afar	0.0037	0.0304	0.123	0.4192	1.010	0.946	1.065
Amhara	0.0021	0.0180	0.117	0.9069	1.002	0.956	1.047
Benishangul-Gumuz	0.0724	0.0265	2.730	0.0063	1.075	1.021	1.132
Dire Dawa	0.0224	0.0333	0.642	0.5206	1.023	0.957	1.090
Gambela	0.0698	0.0322	2.167	0.0302	1.072	1.007	1.142
Harari	0.0318	0.0331	0.960	0.3371	1.032	0.967	1.101
Oromia	0.0407	0.0251	1.617	0.1058	1.042	0.991	1.094
SNNPR	0.0672	0.0268	2.507	0.0121	1.069	1.015	1.127
Somali	0.0956	0.0395	2.420	0.0181	1.100	1.018	1.189
Residence(Ref=urban)							
Rural	0.0129	0.0229	0.565	0.5723	1.013	0.969	1.060
Religion(Ref=Orthodox**)**							
Muslim	0.0616	0.0176	3.497	<.0001	1.063	1.027	1.101
Protestant and other	0.0356	0.0204	1.749	0.0802	1.036	0.996	1.078
Family size	0.0142	0.0031	4.572	<.0001	1.014	1.008	1.020
Wealth (Ref= Lowest**)**							
Middle	-0.0207	0.0101	-2.049	0.0410	0.980	0.949	1.002
			-2.787	0.0050	0.937	0.895	0.981

Highest	-0.0649	0.0232					
Age 1st birth(Ref=<=15)							
16-19	0.1346	0.0142	9.450	<.0001	1.144	1.113	1.177
20-40	0.1193	0.0177	6.733	<.0001	1.127	1.088	1.167
No living children	0.1780	0.0040	44.354	<.0001	1.018	1.185	1.204
Age at 1st sex	-0.0253	0.0022	-11.459	<.0001	0.975	0.971	0.979
Occupation of father (Ref=Didn't work)							
Did have work	0.0683	0.0120	5.692	<.0001	1.071	1.021	1.089
Year of education	-0.0212	0.0020	-10.468	<.0001	0.979	0.975	0.983
No- of visits	0.0153	0.005	2.966	0.0030	1.015	1.005	1.026
Contraceptive use (Ref=No)							
Yes	0.0461	0.0254	1.813	0.0697	1.047	0.996	1.101
Ideal no children	0.0066	0.0022	2.954	<.0001	1.007	1.002	1.011

4.3 Multilevel Count Analysis of the Data

In the multilevel analysis, two -level structure is used with regions as the second-level units and individual mother as the first level units. In this study we consider multilevel models to allow and between-region variance of total number of ever born children. The data have a two -level hierarchical structure with 9602 mothers at level 1, nested within 11 regions states at level 2.

In the multilevel Poisson regression analysis empty model, random intercept model, random coefficient model and random coefficient with interaction were used in the analysis and deviance based chi-square test used for model comparison, but before fit multilevel model we compared (ordinary) single level and multilevel count models.

4.3.1. Test of Heterogeneity

A likelihood ratio test is applied to assess heterogeneity of a total number of ever born children per mother among the 11 regions. Comparisons of multilevel (Poisson and NB) models with their single level count model, with LRT statistic given in Table 4.7. The values of LRT's for each model is larger than the critical value or with p-value < 0.05. Thus, there is an evidence of heterogeneity of total ever born child across regions and also observed that multilevel count regression model is best fit over the ordinary (single level) count regression models (Table 4.7).

Table 4.7. Likelihood ratio test for single level and multilevel count models

Criterion	Model	Value	p- value
LRT	Poisson	18.202	0.0336
	NB	172.989	0.0001

4.3.2 Model Selection Criteria

Table 4.8. Shows that deviance, AIC and BIC for model selection and fit criteria. A lower value of these criteria suggests a better fit. Since multilevel NB regression model has smaller value in deviance, AIC and BIC. Consequently, we conclude that in this study multilevel NB regression model is better than the multilevel poison model. In overall, all criteria shows that the multilevel NB model better fit for number of total children ever born than other models thus are (ordinary Poisson and NB and multilevel NB) models.

Table 4.8. Multilevel negative binomial regression model compassions

Model	**selection criteria**		
	Deviance	***AIC***	***BIC***
Intercept only	44342	44348	44349
Random intercept	30317	30454	30583
Random slope/coefficient	30452	30458	30587

4.3.3. Parameter estimates of multilevel NB regression model

The results from table 4.9 showed that the estimates random effect for count part of the model for total number of ever born children per mother varies among regions. Since, the random intercept is statistically significant at 5% level of confidence. Also from the table we showed that to interring all covariates regional variations decreased from 0.04891 (level-two variance without covariates) to 0.0071267, it indicates that there is a significant variation between regions in total number of ever born children per mother.

Results of Multilevel NB regression model; wealth index, year of education, contraceptive use and age of mother at 1st sex are negatively associated with total number of ever born children. However, family size, age of mother, age of mother at first birth, number of living child's, father's occupation, region, ideal number of children and number of visits are positively associated with a total number of ever born children.

Religion of respondents has significant positive impact on the number of ever born child per mother. Implies that the expected number of ever born child for Muslim believer women's are higher than 1.066 times as compared orthodox believer's women's. Similarly the expected number of ever born child for protestant and other believer women's are higher than 1.072 times as compared orthodox believer's women's.

The findings of this study also showed that, current age of mother is a significant positive impact on number of total children ever born per mother. For a unit increased each year in age of mother, then the expected number of total children ever born per mother is increased by1.014 times.

And also number of living children of mothers has significant positive impact on number of total children ever born per mother. Particularly, for a unit increase in living child of mother, the expected number of total children ever born is increased by 1.196 times.

A similar way age of mother at first sex has significant negative impact on total number of children ever born per mother. Particularly for one unit increased in age at first sex of mother the expected number of total children ever born decreased by 0.976 times. Also number of visits of mother has significant positive impact on total number of children ever born per mother. Particularly for one unit increased in number of visits the expected number of total children ever born increased by 1.016 times. Also ideal number children has significant positive impact on total number of children ever born per mother. Particularly for one unit increased in ideal number child the expected number of total children ever born increased by 1.007 times. Similarly a family size of household's has positive significant effect on total number of children ever born per mother. Particularly for every unit increased in family size the expected number of total children ever born increased by 1.014 times.

This study shown that age at first birth has a significant positive impact with total number of ever born children per mother. The expected number of ever born children with age at first birth lies between 16 and 19 is higher than 1.142 times as compared to age at first birth less than or equal to 15. Similarly The expected number of ever born children with age at first birth lies between 20 and 40 is higher than 1.196 times as compared to age at first birth less than or equal to 15.

Contraceptive use of the mother has significant negative impact on the number of ever born child per mother. Implies that the expected number of ever born child for using Contraceptive are lower than 0.9557 times as compared not using Contraceptive. Similarly, wealth index of the respondents has significant negative impact on the number of ever born child per mother. Implies that the expected number of ever born child for highest wealth index are lower than 0.939 times as compared lowest wealth index. Also occupation of father has significant positive impact on the number of ever born child per mother. Means the expected number of ever born child of farther who has job attachment is higher than 1.050 times that of father who has job attachment.

Table 4.9. Parameter estimates of multilevel NB regression model

	Estimation	S.E	Z-value	p-value	IRR	95%CI for IRR	
						Lower	upper
Current age of mother	0.0217	0.0010	22.30	<.0001	1.0220	1.0200	1.0239
Residence(Ref=urban)							
Rural	0.0182	0.0221	0.82	0.410	1.0184	0.9752	1.0636
Religion(Ref=Orthodox**)**							
Muslim	0.0641	0.0148	4.99	<.0001	1.0662	1.0460	1.1085
Protestant and other	0.0695	0.0210	3.32	<.0001	1.0720	1.0287	1.1170
Family size	0.0136	0.0031	4.38	<.0001	1.0137	1.0075	1.0199
Wealth (Ref= Lowest**)**							
Middle	-0.0198	0.0131	-1.50	0.1323	0.9804	0.9555	1.0060
Highest	-0.0625	0.0226	-2.76	0.0058	0.9394	0.8986	0.9821
Age 1st birth(Ref=<=15)							
16-19	0.1329	0.0142	9.33	<.0001	1.1422	1.1107	1.1745
20-40	0.1180	0.0177	6.67	<.0001	1.1253	1.0869	1.1649
No living children	0.1792	0.0040	44.50	<.0001	1.1963	1.1869	1.2058
Age at 1st sex	-0.0247	0.0022	-11.20	<.0001	0.9756	0.9714	0.9798
Occupation father (Ref=Didn't work) Did have work	0.0492	0.0163	3.02	<.0001	1.0505	1.0174	1.0846
Year of education	-0.0213	0.0020	-10.71	<.0001	0.9789	0.9751	0.9827
No- of visits	0.0156	0.0051	3.04	0.0023	1.0158	1.0056	1.0261
Contra _use (Ref=No) Yes	-0.0453	0.0150	-3.02	0.0025	0.9557	0.9280	0.9842
Ideal no children	0.0073	0.0022	3.29	<.0001	1.0074	1.0030	1.0118
/lnalpha	-1.1056	0.0543	-20.43	<.0001	0.3310	-1.2121	-0.9990
Intercept ($\sigma^2 u_0$)	0.0113	0.0027	4.13	<.0001	1.0113	1.0016	1.0837

Key: Significant (P-value< 0.05). Ref: Reference category

4.4 Discussion of the Results

Fertility is one of the elements in population dynamics that has significant contribution towards changing population size and structure over time in the world. High fertility hinders achieving national goals such as food sufficiency, universal primary education, improving the accessibility of health care services to the largest possible number in the shortest possible time and employment and housing conditions are remarkably difficult, and mothers with higher parities have greater risk of death related to pregnancy and child birth and children from large family size have high risk of mortality. In general there is unbalance the number of children ever born and resource between the regions. Therefore, this study was to identify the determinants of fertility statues of Married women in Ethiopia using Ethiopian demographic and health survey (EDHS 2016) data using count regression models.

According to the results, wealth index is an important socio-economic predictor on total number of ever born children; that is, fertility rate decreases with increase in wealth index. This result in shows with the previous study that, there is reverse relationship between income and fertility (Priya and Kshatriya, 2016), (Muluneh, 2016), (Dana, 2018).

According to the results, contraceptive use is an important socio-economic predictor on total number of ever born children; that is, fertility rate decreases with increase in contraceptive use. This result in shows with the previous study that, there is reverse relationship between contraceptive use and fertility(Muluneh, 2016).

Further, the result of this study indicated that religion of respondent has significantly associated with ever born children with those who has believer of Muslim and those who has believer of protestant and others religion having higher chances of having more children as compared to those who has believer of Orthodox religions. This is consistent with the study of (Priya and Kshatriya, 2016), (Ayele, 2015)

Also ideal number children has significant positive association on total number of ever born children; that is fertility rate increased as ideal number of children increased similarly from the previous study says that, ideal number of children showed a significant positive effect on fertility (Priya and Kshatriya, 2016),(Awad and Yussof, 2017).

The result of this study said that, number of visits has significant positive impact on total number of children ever born per mothers, implies that fertility rate increased as number visit increased similarly from the previous study by (Dana, 2018). Also this study said that for every unit increase in age of mother, total number of ever born children increases and this is similar to the findings (Muluneh, 2016).

In addition to this, the result showed that age of mother at first birth are positively associated with total children ever born. It was also found in this study that women who got their first child at earlier ages were more likely to have more children than those women who got at later ages (Muluneh, 2016).

In addition to this the study said that, age of mother at first sex has significant negative impact on total number of children ever born per mother. Implies for every unit increment in age at first sex of mother the fertility rate decreased, this study similar with (Berlie and Alamerew, 2018),(Ayele, 2015).

This study showed that women's year of educational attachment were significant negative relationship with total children ever born per women. Since, age of women was loss/miss much time by learning and also we expect educated mothers were can facilitates family planning like use contraceptive and to predict future plan for their children (Dwivedi et al., 2016),(Cesar Augusto Oviedo Tejada1, 2017)

Also result of this study shows that occupation of father has significant positive impact on the number of ever born child per mother, means that father who has job attachment were more likely to have more children than those father who has not job attachment (Pandey and Kaur, 2015).

This study showed that family size has significant positive association on total number of ever born children; that is fertility rate increased as for a unit increments of family size. But as family size in the household increases, fertility rate decreased (Ayele, 2015).

In this study, two count regression models were considered. The variance of total number of ever born children was larger than its mean, suggesting the possibility of over-dispersion. In addition, the over-dispersion parameter alpha was found to be significantly the different from zero in NB regression model, in addition to this model compression to be supported this idea due to this reason ordinary NB regression model is better fit than single level Poisson regression model.

In multilevel count regression analysis, individual mothers are considered as lower level nested within the various regions in Ethiopia. Before going to in the multilevel approach first compare likelihood ratio test of singe level and multi-level models. To observe test of heterogeneity across the region, implies that if there are differences in number ever born children between regions. The test suggested that, the number of ever born children varies between regions and multilevel count model better fit than the single level count model. Among the two multilevel count regression model, multilevel NB model is the best model for the heterogeneity of the number of ever born children per mother among regional level in Ethiopia.

CHAPTER FIVE

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The purpose of this study has identified socioeconomic demographic, health and environmental related determinants and assessed regional variation of the total number of ever born children per mother at the reproductive age of (15-49) have in their lifetime using count regression model. Data from EDHS 2016 were used for analysis. In this study, two count regression models were considered. The variance of total number of ever born children was larger than its mean, suggesting the possibility of over-dispersion. In addition, the over-dispersion parameter alpha was found to be significantly different from zero in NB regression model.

In this study ordinary or single level and multilevel count models to be considered. From the result of single level regression mode, ordinary NB regression model is better fit than single level Poisson regression model. The result of ordinary NB regression model showed that wealth index, year of education and age of mother at 1st sex, family size, age of mother, age of mother at first birth, number of living child's, father's occupation, region, ideal number of children and number of visits are statistically significant factors on total number of ever born children per mother in Ethiopia.

In multilevel count regression analysis, individual mothers are considered as lower level nested within the various regions in Ethiopia. Before going to in the multilevel approach first compare likelihood ratio test of singe level and multi-level models. To observe test of heterogeneity across the region, implies that if there are differences in number ever born children between regions. The test suggested that, the number of ever born children varies between regions and multilevel count model better fit than the single level count model. Among the two multilevel count regression model, multilevel NB model is the best model for the heterogeneity of the number of ever born children per mother among regional level in Ethiopia.

Generally result shows that Multilevel NB regression model was found to be the most appropriated and preferred model than the single level and Multilevel Poisson regression model by supported test and different model comparisons including AIC, BIC and deviance.

From the result of multilevel NB regression model, the Random Intercept model was found to be the best fit for total number of ever born children per mother. Variables such as wealth index, year of education, age of mother at 1st sex, family size, age of mother, age of mother at first birth, number of living child's, father's occupation, region, ideal number of children and number of visits have statistically significant effect on a total number of ever born children.

5.2 Recommendations

- Regions like Somali, Benishangul-Gumuz, SNNP and Gambela should be given special attention to reduce fertility rate through empowering women and campaigns for the further increasing year of education, age of mother at first sex, contraceptive use and wealth index of households.
- In addition to this the end user governmental and non-governmental organizations could take intervention measures and set appropriate plans to tackle fertility problems like to the reduce number of children ever born per women in order to balance resource between the regions, also to be useful for policy making, monitoring and evaluation different activities
- Further researchers in the area should incorporate other important missed variables and also focus on those areas having high fertility in order to explore the problem and to forward a solution.

REFERENCE

ABEBE, Y., KONDALE, M., JILO, G. K., HEBO, S. H. & SIDAMO, N. B. 2018. DETERMINANTS OF HIGH FERTILITY AMONG MARRIED WOMEN IN ANGACHA DISTRICT, KAMBETA TEMBERO ZONE, SOUTHERN ETHIOPIA. Ethiopian Journal of Reproductive Health, 10.

ADUGNA, A. 2014. Population policy and projection.

AGENCY, C. S. 2007. Population and Housing Census. Statistical Report: central statistical agency.Addis Ababa, Ethiopia.).

ALENE, G. D. & WORKU, A. 2009. Estimation of the total fertility rates and proximate determinants of fertility in North and South Gondar zones, Northwest Ethiopia: An application of the Bongaarts' model. Ethiopian Journal of Health Development, 23.

*ATSBAHA, G., HAILU, D., BERHE, H., SLASSIE, A. & YEMANE, D. 2016. Determinants of High Fertility among Ever Married Women in Enderta District, Tigray Region, Northern Ethiopia. J Health Med Informat, 7**,** 243-248.*

*AWAD, A. & YUSSOF, I. 2017. Factors Affecting Fertility–New Evidence from Malaysia. Bulletin of Geography. Socio-economic Series, 36**,** 7-20.*

*AYALEW, J., MOGES, H. & WORKU, A. 2014. Identifying factors related to the survival of AIDS patients under the follow-up of antiretroviral therapy (ART): The case of South Wollo. Int. J. Data Envelopment Anal. Oper. Res, 1**,** 21-27.*

*AYELE, D. G. 2015. Determinants of fertility in Ethiopia. African health sciences, 15**,** 546-551.*

BANK, W. 2009. World development indicators.Washington DC: World Bank.

BERLIE, A. B. & ALAMEREW, Y. T. 2018. Determinants of Fertility Rate among Reproductive Age Women (15-49) in Gonji-Kollela District of the Amhara National Regional State, Ethiopia. Ethiopian Journal of Health Development, 32.

BOCK, R. D. & MISLEVY, R. J. 1989. A hierarchical item response model for educational testing. Multilevel analysis of educational data. Elsevier.

*BONGAARTS, J. 2008. Fertility transitions in developing countries: Progress or stagnation? Studies in family planning, 39**,** 105-110.*

BUREAU, P. R. 2015. World population Data sheet with a special focus on women's empowerment. Washington, DC, Population Reference Bureau. .

*CAMERON, A. C. & TRIVEDI, P. K. 1986. Econometric models based on count data. Comparisons and applications of some estimators and tests. Journal of applied econometrics, 1***,** *29-53.*

CAMERON, A. C. & TRIVEDI, P. K. 2013. Regression analysis of count data, Cambridge university press.

CARLIN, B. P., GELFAND, A. E. & BANERJEE, S. 2014. Hierarchical modeling and analysis for spatial data, Chapman and Hall/CRC.

CESAR AUGUSTO OVIEDO TEJADA1, L. V. M. T., FLA ´VIA KATREIN DA COSTA1,FRANCIELE HELLWIGCESAR AUGUSTO OVIEDO TEJADA1, Lı ´VIA MADEIRA TRIACA1, FLA ´VIA KATREIN DA COSTA1,FRANCIELE HELLWIG 2017. The sociodemographic, behavioral,reproductive, and health factors associated with fertility in Brazil.

*CHAMPION, A. G. 2001. A changing demographic regime and evolving poly centric urban regions: Consequences for the size, composition and distribution of city populations. Urban Studies, 38***,** *657-677.*

CHATFIELD, C., ZIDEK, J. & LINDSEY, J. 2010. An introduction to generalized linear models, Chapman and Hall/CRC.

COALE, A. J. & HOOVER, E. M. 2015. Population growth and economic development, Princeton University Press.

*DANA, D. D. 2018. Binary Logistic Regression Analysis of Identifying Demographic, Socioeconomic, and Cultural Factors that Affect Fertility Among Women of Child bearing Age in Ethiopia. Science Journal of Applied Mathematics and Statistics, 6***,** *65.*

*DWIVEDI, V., SEDIADIE, T. & AMA, N. 2016. Factors affecting children ever born (CEB) in Botswana: application of Poisson regression model. Res J Math Stat Sci, 4***,** *1-9.*

ECONOMICDEVELOPMENT, P. D. M. O. F. A. 2006. Ethiopia: Population Images.

EDHS 2016a. Ethiopia Demographic and Health Survey.

EDHS 2016b. Ethiopian Demographic and Health Survey.

EDHS 2016c. Ethiopian Demographic and Health Survey.

*EYASU, A. M. 2015. Multilevel modeling of determinants of fertility status of married women in Ethiopia. American Journal of Theoretical and Applied Statistics, 4***,** *19-25.*

FADDY, M. & SMITH, D. 2011. Analysis of count data with covariate dependence in both mean and variance. Journal of Applied Statistics, 38, 2683-2694.

FORUM, W. E. 2012. Global agenda council on population growth
2012-2014. .

GOLDSTEIN, H. 1995. Hierarchical data modeling in the social sciences. Journal of Educational and Behavioral Statistics, 20, 201-204.

GOLDSTEIN, H. 2011. Multilevel statistical models, John Wiley & Sons.

HARRIS, T., HILBE, J. M. & HARDIN, J. W. 2014. Modeling count data with generalized distributions. Stata Journal, 14, 562-579.

HILBE, J. M. 2011. Negative binomial regression, Cambridge University Press.

HOX, J. J., MOERBEEK, M. & VAN DE SCHOOT, R. 2017. Multilevel analysis: Techniques and applications, Routledge.

IDDI, S. & MOLENBERGHS, G. 2012. A combined overdispersed and marginalized multilevel model. Computational Statistics & Data Analysis, 56, 1944-1951.

LAWLESS, J. F. 1987. Negative binomial and mixed Poisson regression. Canadian Journal of Statistics, 15, 209-225.

MAAS, C. J. & HOX, J. J. 2002. Robustness of multilevel parameter estimates against small sample sizes. Unpublished Paper, Utrecht University.

MORGAN, A. S., MARLOW, N., DRAPER, E. S., ALFIREVIĆ, Z., HENNESSY, E. M. & COSTELOE, K. 2016. Impact of obstetric interventions on condition at birth in extremely preterm babies: evidence from a national cohort study. BMC pregnancy and childbirth, 16, 390.

MULUNEH, E. K., ZEWOTIR, TEMESGEN, & YAYEH, TILAYE MATEBE 2016. Fertility Status of Married Women and Its Determinants in Ethiopia. .

NATIONS, U. 1993. population division Total fertility (children per woman) Medium variant (1950-2015).

ORGANIZATION, W. H. 2005. Make every mother and child count.Switzerland.

PANDEY, R. & KAUR, C. 2015. Modelling fertility: an application of count regression models. Chinese Journal of Population Resources and Environment, 13, 349-357.

PRIYA, R. & KSHATRIYA, G. 2016. A study of factors affecting fertility among Dhangars of Madhubani District.

SCHALL, R. 1991. Estimation in generalized linear models with random effects. Biometrika, 78, 719-727.

SELLERS, K. F. & SHMUELI, G. 2013. Data dispersion: now you see it… now you don't. Communications in Statistics-Theory and Methods, 42, 3134-3147.

SHARMA, M. P. 2015. The Determinants of Fertility among Women of Reproductive Age in Nepal. The Journal of Development and Administrative Studies (JODAS), Vol. 23(1-2), pp. 55-68.

SHERAZI, M. H. & BUKHARI, U. 2019. Common Gynecology and Obstetric Symptoms for the Objective Structured Clinical Examination. The Objective Structured Clinical Examination Review, 323.

UN 2009. Department Economic and Social Affairs (UNDESA) (2009) World Fertility Patterns: New York.

UN 2015a. World Fertility Report, Department of Economic and Social Affairs.

UN, U. 2015b. The millennium development goals report. UN New York.

VER HOEF, J. M. & BOVENG, P. L. 2007. Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? Ecology, 88, 2766-2772.

APPENDIX

No	Variable	Description	Code
Demographic factors			
1	Sex	Sex of household head	0=Female 1=Male
2	Age	Age of mother	Continuous variable
3	Age at 1st birth	Age of mother at the 1st birth	0=<=15 1=16-19 2=20-40
4	Age at 1st sex	Age of mother at the 1st sex	Continuous variable
5	Family size	Number of household members	Count variable
6	Living children	Total number of living children	Count variable
7	Age of father	Age of father	Continuous variable
Environmental factors			
8	Region	Region of the household	0=Tigray 1= Afar 2= Amhara 3= Oromia 4= Somali 5=Benishangul-Gumuz 6= SNNP 7= Gambela 8= Harari 9= Addis Abeba 10= Dire Dawa
9	Place of residence	Type of place of residence	0= Urban 1= Rural
Socio-economic factors			
10	Mother's Education	Mother's education level	0 = No education 1 = primary 2 = secondary and above
11	Father's education	Father's education level	0 = No education 1 = primary 2 = secondary and above
12	Religion	Religion of house hold	0=Orthodox 1=Muslim 2=protestant and Others
13	Mother's occupation	Occupation of mother	0=No working 1=had working
14	Father's occupation	Occupation of father	0=No Working 1=Had Working

15	Wealth index	Wealth index of household's	0=Lowest 1=Medium 2=Highest
16	Media Exposure	Household ever listened to any mass media	0=No 1=Yes
17	Ideal numbers	Ideal number of children	Count variable
18	Recent sexual activity	Recent sexual activity of Women	0= Not active 1= Active
19	Year of education	Live with education in single years	Count variable
Health related factors			
20	Knowledge of contraceptive use	Knowledge of contraceptive use	0=No 1=Yes
21	Contraceptive use	Women' are using Contraceptive	0=No 1=Yes
22	Number of visits	Number of visits	Count variable

Model comparison for single level count model results

	Criteria			
Model	**Full Log Likelihood**	**AIC**	**AICC**	**BIC**
Poisson	-15750	31553	31553	31747
Negative binomial	-15245	30546	30546	30747

Model comparison for multilevel count model results for intercept only model

	Criteria			
Model	**Deviance**	**AIC**	**AICC**	**BIC**
Poisson	46746	46750	46750	46751
Negative binomial	44342	44348	44348	44349

Fixed Effects of the intercept only model for NB count model

	Estimation	S.E	Z-value	p-value	IRR	95%CI for IRR Lower	upper
Intercept	1.2685	0.0672	18.8877	<.0001	3.555	3.12	4.07
Intercept ($\sigma^2 u_0$)	0.0489	0.0213	2.2995		1.050	1.01	1.09

Multilevel Poisson random intercept model

Total_children		Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Age Mother		.0217	.0010	22.30	0.000	.0198	.0236
Residence	Rural	.0182	.0221	0.82	0.410	-.0251	.0616
Religion	Muslin	.0740	.0148	4.99	0.000	.0449	.1031
	Others	.0695	.0210	3.32	0.001	.0284	.1106
Family Size		.0136	.0031	4.38	0.000	.0075	.0197
Wealth index	Middle	-.0198	.0131	-1.51	0.132	-.0455	.0059
	Highest	-.0625	.0226	-2.76	0.006	-.1069	-.0181
Age_1birth	16-19	.1329	.0142	9.33	0.000	.1050	.1608
	20-40	.1180	.0177	6.67	0.000	.0833	.1527
Living child		.1792	.0040	44.91	0.000	.1714	.1871
Age_1sex		-.0247	.0022	-11.20	0.000	-.0290	-.0204
Ocu_Father	have work	.0492	.0163	3.02	0.003	.0172	.0812
No visits		.0156	.0051	3.04	0.002	.0056	.0257
Year Edu		-.0213	.0020	-10.71	0.000	-.0253	-.0174
Ideal		.0073	.0022	3.29	0.001	.0030	.0117

Knowledge_CU	Yes	.0453	.0250	1.82	0.069	-.0036	.0942
_cons		-.0381	.0597	-0.64	0.524	-.1551	.0790

------------------------+---

Region

var(_cons)	.0001	.0002519	2.59e-06	.0062159

Log likelihood = -15759.767 Prob > chi2 = 0.0000