ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAMME

DEPARTMENT OF STATISTICS

Factors Associated with number of Children that women in Ethiopia would like to have in their lifetime (Application of Count Regression Model)

Daniel Nigatu

Advisor: Mekonnen Tadesse (Associate Prof.)

A Thesis Submitted to the Graduate programs of Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics (Applied Statistics)

June, 2017

Addis Ababa University

Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY
GRADUATE STUDIES PROGRAMME
DEPARTMENT OF STATISTICS

Factors Associated with number of Children that women in Ethiopia would like to have in their lifetime

Daniel Nigatu

A Thesis Submitted to the Graduate programs of Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics (Applied Statistics)

Advisor: Mekonnen Tadesse (Associate Prof.)

June, 2017
Addis Ababa University
Addis Ababa, Ethiopia

Addis Ababa University
School of Graduate Studies

This is to certify that the thesis prepared by Daniel Nigatu, entitled: Count Regression models for Factors associated with number of Children that women in Ethiopia would like to have in their lifetime and submitted in partial fulfillment of the requirements for the Degree Master of Science in Statistics complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Name	Signature	Date
Advisor: Mekonnen T. (Assoc. Prof.)	___________	__________
Examiner: Birhanu Teshome (PhD)	___________	_________
Examiner: Dejen Tesfaw (PhD)	____________	_________

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all sources of materials used for the thesis have been duly acknowledged.

Name: Daniel Nigatu

Signature: ___________________________

Place of submission: Department of statistics, College of Natural and computational sciences, Addis Ababa University

Date of submission: May 2017

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

Mekonnen Tadesse (Assoc. Prof.) ______________________________

Advisor's Name **signature**

Acknowledgment

First of all I wish to express my profound gratitude to the **Almighty God for his love**, grace and protection throughout the course of my study in this university.

Secondly, Special thanks must go to my advisor, **Mekonnen Tadesse (Assoc.Prof.),** for his supervision, kind help, invaluable advice, professional guidance, useful comments and suggestions throughout my study. It is explicable that without his closer follow-up and continuous encouragement with valuable comments this paper would not be able to produce this quality Paper.

Thirdly, my deepest thanks go to all members at Addis Ababa University especially, my instructors and all individuals in the Department of Statistics for their moral support and positive attitude when I needed them during my course work, and research period.

My heartfelt appreciation goes to **Dr. Gemoraw Addinew (Econometrician and Consultant Manager of DMISS PLC.)** for a great number of things: for his inspiration, for his encouragement, for sharing his knowledge, for forcing me to dig deeper into my research for his timely assistance and helping me to find the solutions.

I am particularly indebted to **Mr. Mario G. (Technical Consultant for World Bank)** for his support and encouragement during the study time by providing technical guidance in statistical software and valuable comments and suggestions.

My special thanks to my family and staff members of **NSDQSCD** especially **Mr**. **Essayas Muleta (NSDQSCD Director)** and also all Sampling Experts for covering my working time, support and encourage me during this study.

Finally, I would like to express my appreciation to all friends who helped in diverse ways and also my classmates who were the source of strength towards the successful completion of the study. I will forever be indebted to you all.

Table of Content

List of Tables

List of Figures

List of Abbreviations

AIC	Akaike Information Criteria
BIC	Bayesian Information Criteria
CEB	Children ever born
CSA	Central Statistical Agency
EA	Enumeration Area
EDHS	Ethiopian Demographic and Health Survey
GLM	General Linear Regression Model
ICF	International Consultancy Fund
INC	Ideal number of children
LL	Log Likelihood
LRT	Likelihood Ratio Tests
MINC	Mean of Ideal number of children
NB	Negative Binomial
NBRM	Negative Binomial Regression Model
NFHS	National Family Health Survey
OLS	Ordinary Least Square
PHC	Population and Housing Census
PMF	Probability Mass Function
PRM	Poisson Regression Model
SNNP	South Nation Nationality of people
TFR	Total Fertility Rate
USA	United States of America
USAID	United States Agency for International Development
V	Vuong
ZI	Zero-Inflated
ZIP	Zero-Inflated Poisson
ZINB	Zero-Inflated Negative Binomial

Abstract

Factors associated with number of Children that women in Ethiopia would like to have in their lifetime

Daniel Nigatu
Addis Ababa University, 2017

Most frequently, fertility preference is defined as desired family size, ideal number of children, and the desire for additional children or the fertility intention. Despite the various strategies and policies to reduce family size, Ethiopia's total fertility rate still remains high at 4.8 partly attributed to strong preferences for large family sizes. The main aim of this study was to explore the factors that influence desired number of children among women in Ethiopia based on data from Ethiopian Demographic and Health Survey conducted in 2011. Count regression models were used to model the desired number of children among women in Ethiopia. Separate analyses were made using the data with sampling weights and without sampling weights and the results revealed that both the descriptive and count regression analyses results using the data with sample weights are different from the results obtained using data without sample weights. Among the count models considered, the Zero Inflated Poisson (ZIP) model was found to be the most appropriate model for analyzing the data on the desired number of children by women in Ethiopia. The ZIP model fit results indicated that religion, wealth index, women age, women education, Place of residence, whether a woman has living children, contraceptive use and women occupation are significantly associated with the desired number of children that women in Ethiopia would like to have in their life time.

Key words: *Ethiopia, fertility rate, Ideal number of children, socio-demographic factors, statistical weight, Count Regression models.*

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. 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; Casterline, 2001). About eight percent of the world's population lives in "high-fertility" countries that have experienced only limited fertility decline to date. In these countries the average woman has five or more children over her lifetime. Most of these countries are in sub-Saharan Africa (UN, 2015).

Ethiopia is one of the developing countries with high fertility and rapid population growth rate. The country's population in 2016 was estimated around 100 million (CSA, 2016), placing it as the second most populous nation in sub-Saharan Africa. According to EDHS 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 2011. Overall, the total wanted fertility rate (TWFR) in Ethiopia is 3.0 children per woman, 1.8 fewer than the total fertility rate (TFR) of 4.8. This suggests that the TFR is 60 percent higher than it would be if unwanted births were avoided (CSA and ORC Macro, 2011). All the efforts to control higher population growth become useless when the household desire larger number of children (Khan and Bari, 2014).

Early age at first marriage, desire for more children and extremely low contraceptive use are some of the major reasons behind such high fertility rate (Assefa, 2001; Kinfu, 2002; Gibson and Mace, 2002).

Ethiopia, being one of the developing countries where subsistence agriculture is the major economic activity, families often prefer large number of children since they are considered as an economic asset rather than liabilities. In rural areas, parents want to have large number of children to get assistance in farming activities (Bairagi, 2001) and emotional as well as economic support during old ages (Fapohunda and Todaro, 2000).

In traditional societies, children are also expected to strengthen the extent of kin relations, which implies not only economic benefits but also physical protection. Like many countries in sub-Saharan Africa, traditional norms and values in Ethiopia are in favour of high fertility. Having many children is considered as a virtue and respect of God in a number of Ethiopian rural communities (Desta and Seyoum, 1998).

The Ethiopian government has been making several efforts to reduce fertility levels since 1993, the first time an explicit national population policy aimed at reducing total fertility rate from 7.7 children per woman to 4.0 by 2015 was launched (NPO, 1993). Increasing age at first marriage to at least 18 years, enhancing women's status through providing them with better employment and educational opportunities, expanding family planning services and information, communication and education on ways and means of limiting family size are some of the strategies designed to implement the population program (Assefa, 2001).

Various international agencies and non-governmental organizations have also made efforts to provide technical, material and financial support to the Ethiopian Government in implementing the population program so that fertility would decline over time. Even though fertility has shown a declining trend at the national level, the transition has not begun in some of the regions. There are clear regional differences in fertility levels and trends in Ethiopia.

Some of the regions like Addis Ababa have achieved a below replacement level of fertility (TFR=1.8) while others such as Oromia (6.2), Somali (6.0), and SNNP (5.6) regions have total fertility rates that are above the national average (EDHS 2011). Like other developing countries, significant variation in fertility level was observed among rural and urban residents of Ethiopia. For instance, according to the 2011 Ethiopian Demographic and Health Survey report, fertility is lower (TFR =2.6 children per woman) in urban Ethiopia than in rural Ethiopia (5.5 children per woman).

Modernization factors such as better access to education and media, employment opportunities in the modern economic sector and wider access to family planning services are some of the major factors that put urban fertility down in the Ethiopian context (Kinfu, 2001; Sibanda et al., 2001). The effect of poverty on fertility decline in major towns of Ethiopia is also not negligible (Eshetu and Mace, 2001).

Fertility preference is defined as desired family size, ideal number of children, and desire for additional children or fertility intentions. The measurements have been used to describe and/or estimate the number of children that people actually want to have.

Fertility preferences are the indicators of general attitudes and possible future course of fertility. Also family planning approval is strongly dependent on fertility preferences. Measuring fertility intentions, and determining the extent to which they predict fertility behavior, is important for population policy and the implementation of family planning programs.

A large proportion of Ethiopians, regardless of their number of living children, consider four or more children to be ideal. Women desired family size of 4.3 children, and men prefer 4.8 children. Women's ideal family size has declined in the last ten years, from 5.3 children in 2000 to 4.5 children in 2005 and 4.3 children in 2011. (CSA and ICF International, 2012).

Fertility desires and intentions are innermost in theoretical and empirical approaches to studying childbearing behavior and it is one of the most significant determinants for future population structure of a country. Fertility trends can be predicted and population growth can be controlled by recognizing the factors that affect fertility preferences and desires. In the setting of the above circumstances, this study attempts to understand fertility behaviors and identify factors associated with the desired number of children among women in Ethiopia.

1.2. Statement of the Problem

The specific nature of the determinant of fertility is intricate and complex. While, fertility behavior influences population growth, it has consequences on resources, employment situation, health and other social facilities. The mechanism of factors affecting fertility is that intermediate variables influence fertility directly, while socio-economic and demographics variables affect fertility indirectly through intermediate variables (Bongaa rts 1978). Some of these factors could be literacy status, occupation, religion, child birth- death ratio, wealth status, place of residence, household headship, contraceptive use, region, reproductive life span, age at first marriage and desired number of children (Behrman, J. R. and Wolfe, B. L.(1984,), (Angeles, 2008).

Most studies have concentrated on family planning as a general measure for fertility preference as much as it has its flaws. Collecting information on family planning as a measure of fertility preference can be relatively complex.

Often it is difficult to get objective responses as questions on family planning are hypothetical in nature. Respondents, especially those illiterate or with little education may find it difficult to understand these questions(Zhang, 2007).this study uses preferred desired number of children among women in Ethiopia as the response variable.

Several studies investigated determinants of fertility in Ethiopia using some set of variables and statistical methods such as logistic regression, survival analysis, and linear regression models. Since, the ideal number of children (INC) data is a count data, Poisson Regression Models (PRM) and Negative Binomial Regression Models (NBRM) have been shown to be statistically more appropriate (Poston, 2002).

Conversely, many studies chose to categorize count variables as binary and did use sampling weights that will make the sample representative of the target population. If we treat ideal number of children as binary variable and ignore the sampling weights, we will be losing some information and get misleading results. Thus, this study aimed to assess the ideal number of children per and the potential factors influencing it using appropriate count regression models and sampling weights that will make the sample representative of the target population. The researcher believes that the use of sampling weights will increase representation and the reliability or precision of the statistical results of the study.

Since the fertility level in Ethiopia is 4.8 children per woman (CSA and ORC Macro, 2011) and some policy makers are worried about the high fertility rate, this study seeks to answer the following research questions:

- ✓ Which socio-demographic factors are affecting Ethiopian women desired Children?
- ✓ What is the average desired number of children among Ethiopian women of reproductive age?

1.3. Objectives of the Study

General Objective

The main objective of the study was to identify socio-economic, demographic factors associated with the desired number of children among Ethiopian women.

Specifically, the study aimed at:

- Estimating the average desired number of children among Ethiopian women of at reproductive age
- Identifying factors associated with the desired number of children among women in Ethiopia using sampling weights
- Comparing statistical results based on data with sampling weights and without sampling weights.

1.4. Significance of the Study

It is hoped that findings from this research could be useful in many ways. The findings are believed to be useful for policy making, monitoring and evaluation activities of the government and different concerned agencies.

Bankole and Audam (2011) believed that most spouses respect their fertility preferences, whether in terms of desired number of children or desire for a future birth. As a result, the investigation of the number of desired children has great importance in fertility studies. In addition, the desired number of children has great importance in determining the tendency of people to use contraception (MacDevittet al., 1996).Ideal number of children is an important measure for estimating levels of unwanted or untimely fertility, forecasting fertility, and evaluating unmet need for contraceptives.

This study is expected to contribute its part by filling the information gap concerning the impact of factors that affect desired number of children and fertility in the country. Finally, the study could be a stepping stone for further studies.

The study is also expected to add to the body of knowledge on fertility and reproductive health issues in Ethiopia. In addition, it could assist in policy formulation and programme design appropriate or fortification of existing ones.

To the best of the researchers' knowledge, there are no studies that have used sampling weights in the analysis of EDHS data and hence, this study could be a baseline for other interested in using sampling weights and make their results more reliable and precise.

1.5. Limitation of the Study

Although many factors are affecting the desired number of children, this study is undertaken to explore a few socio-demographic factor that affect fertility in Ethiopia. The following are some of the limitations of our study.

- Some demographic variables like contraceptive awareness and use, abortion, frequency and age at sexual contact initiation might create social desirability bias.
- The data used in this study were from the EDHS 2010/11. Thus, the results may not necessarily reflect the current situation in Ethiopia.

1.6. Organization of the Thesis

This thesis has been organized as follows. In Chapter 1, the background, objectives, scope and limitation of the study are presented. The second Chapter provides the literature review. The data source, methodology and the research hypotheses are given in Chapter three. The fourth Chapter is the analysis, findings and discussion Chapter. Finally, conclusions and recommendations are made in Chapter five.

2. Literature review

2.1. General

Fertility is the natural capability of giving life. Fertility is one of the three major components of population dynamics that determine the size and structure of the population of a country. Differential in fertility behavior and fertility levels in different areas and among population strata or characteristics have been among the most pervasive finding in demography (Ramesh 2010). It is also a complex variable affected by multiple factors. According to Bongaarts (1978), factors affecting fertility are broadly classified into proximate (direct) and distal (indirect) factors. The proximal factors are bio behavioral factors, like being sexually active, use of contraceptive, duration of postpartum in fecund ability, abortion and sterilizing which affect fertility directly, whereas, distal determinant are socio-cultural factors which affect fertility indirectly through affecting the bio-behavioral factors. Human fertility is responsible for the biological replacement and maintenance of the human species. In fact, fertility is a major counteracting force to population attrition from mortality and therefore, has a significant expansionary force in population dynamics. However, the phase of actual reproductive performance is counted in terms of the physiological potential of a woman to conceive and bear children. This phase is termed as the fecund period, which has two extremes, viz., menarche and menopause. In demographic studies, the reproductive span i.e., the child-bearing period of women is usually taken as between 15 to 49 years of age. Thus, a fecund woman may or may not be fertile but a fertile woman must be fecund. The main events or phenomena associated with fertility are age at menarche and age at menopause. It is found that a later age at marriage reduces fertility. Educational level, economic status, religious attitudes, women's work participation etc. are other factors affecting fertility in addition to contraception control practice and attitudes (Samson and Mulugeta 2009).

Variables that influence fertility include: women age (Bongaarts 1978), education (Sharma, 1998), work status of women (Mason and Palan 1981), Place of residence (Abdul Hakim 1994), contraception use (Haile and Enqueslassie 2006), Religion (Caldwell and Caldwell 1987:409), Economic status (Hakim and Miller, 1996).

2.2. Theoretical framework

The economic model of fertility originated from the studies of Becker (1960) and Becker and Lewis (1973), which is known as "Chicago-Columbia‟ model. Becker (1960) introduces the "quantity-quality trade-off" concept of children. He assumes that children are equivalent to consumer durable goods to analysis the demand for children in the household. He also assumed that the preference for children is exogenous, i.e., it is not determined by economic factors. The demand for children depends on women's wage and family income. When women's wages increase, the opportunity cost of having children increases and the demand for children declines. When family income increases, the effect on the demand for children depends on the relative strength of the income and the substitution effect. Households could increase both the number and the quality of children raises the cost of raising children which decreases fertility, which is known as the substitution effect. Becker (1960) states that higher family income results in fewer children of higher quality when the substitution effect is stronger than the income effect.

A crucial assumption of the "Chicago Columbia‟ model is that households can separate quality and quantity of children in their decision making process. Becker (1960) also assumes that in developed countries, the income elasticity of quantity of children is positive but small. However, the income elasticity of quality of children is relatively large due to social pressure, as rich (or poor) families have to maintain the quality of their children according to their status.

He mentioned that families with excess children cannot afford to increase the income elasticity of quality of children. With an increased use of contraceptive, family can increase income elasticity of quality of children by reducing quantity (supply) of children. Moreover, Becker (1976) states that contraceptive knowledge also affects the demand for children. Thus, unequal distribution of contraceptive knowledge across regions may affect the demand for children.

According to Easterlin (1966), if expected income is higher than actual income, then the household's actual number of children will be lower than their expected number of children. He states that the couples "choice of the expected number of children depends on the parents" childhood experience. For instance, wives and husbands who grew up in large families tend to have a higher number child.

Easterlin et al. (1980) also emphasize the rule of biological or supply side factors, such as nutrition and health condition of women, age at women first cohabitation, in fertility determination.

2.3. Empirical Studies

The studies by Behrman and Wolfe (1984), Ainsworth et al. (1996), Osili and Long (2008), Kabeer (2001), kabir et al. (2001), and Lam and Duryea (1999) employed a linear model (OLS) to identify the determinants of women fertility.

Behrman and Wolfe (1984) attempted to relate both "Chicago-Columbia" and "Pennsylvania" models using data on women who completed their fertility. They collected data for women in Nicaragua in 1977-78 and used the number of living children to a woman as the dependent variable. Women's education and household income were used as the "Chicago-Columbia model" variables. They included women health status, women age, age at first marriage, age at first cohabitation, average length of breast feed, and average calorie intake as biological supply variables of the "Pennsylvania" model.

They also included type of marriage, number of sibling, birth rate, childhood residence as expectation building variables of “Pennsylvania‟ model. They find significant determinants in both “Chicago-Columbia‟ and “Pennsylvania‟ variables. They mention that since the “Chicago-Columbia‟ model doesn’t emphasize supply side variables, taste, and other factors for estimating fertility, models based uniquely on the “Chicago-Columbia‟ model may over-estimate the effect of women’s education and household income, leading sometimes to misleading results.

Studies by Ainsworth et al. (1996), Osili and Long (2008) investigated whether women’s schooling negatively affects fertility in African countries. Contrary, to the previous study, these studies use children ever born for each woman as the dependent variable in their models. Ainsworth et al. (1996) estimated fertility of women of fourteen sub-Saharan countries.

They find that women’s education is negatively related to fertility in thirteen sub-Saharan countries, the only exception being Senegal. Using the 1999 Demographic and Health Survey from Nigeria, Osili and Long (2008) investigated whether the negative association between fertility and education was caused by the introduction of universal primary education.

Their assumption was that schooling is endogenous to fertility determination, since fertility choice disrupts schooling. They used exposure to universal program as an instrumental variable (IV) for years of schooling of women.

Using difference in difference method, they estimated the model in both OLS (without instrument) and IV variable approach. In both cases, they did find that the coefficient estimates are negative for women education, but the IV estimates are higher than OLS ones.

Kabeer (2001) and Kabir et al. (2001) also used children ever born as a measure of women's fertility. Kabeer (2001) estimated separate model for different age groups (i.e., 12 to 19, 20 to 40, 40 above) using 1989 Bangladesh Fertility Survey. She concluded that for all age group, men and women's education, wealth, and working status of women are negatively related to fertility and that Muslim women had higher fertility compared to other religion follower women. She also found rural-urban difference for age group 20 to 40, and 40 above, but not for 12 to 19. Kabir et al. (2001) used data set from 1989 Bangladesh fertility survey, Bangladesh Demographic and Health Survey 1993-1994, and 1996-1997. They estimated fertility determinants separately for each data set. They found out that women fertility is negatively related to women's education, employment, place of residence, and access to mass media.

A study in China showed that the preference for a small family was associated with younger age, urban residence, and higher level of education (Ding and Hesketh, 2006). A similar study indicated that men and women with low levels of education were likely to have high mean numbers of children (NSF 2006). Dommaraju and Agadjanian (2009) explained that changes in fertility regime in Bangladesh in most cases are not due to changes in women's status, but due to changes in the reproductive behavior of illiterate women. In most of the studies, which considered the relationship between female education and fertility in Bangladesh, a significant, linear and inverse relationship was found (Akmam 2002).

Bongaarts (2010) in his study among 30 sub-Saharan countries to analyze the causes of educational differences in fertility found out that women with secondary or higher education have on average lower fertility than women with no education (3.4 vs. 6.3 births per woman), which is also the case in desired family size (3.7 vs. 5.6 births per woman). Additionally, there are differences by level of education in the relationships between reproductive indicators. As education rises, fertility is lower at a given level of contraceptive use, contraceptive use is higher at a given level of demand and demand is higher at a given level of desired family size. The most plausible explanations for these shifting relationships are that better-educated women marry later and less often, use contraception more effectively, have more knowledge about and access to contraception, have greater autonomy in reproductive decision making, and are more motivated to implement demand because of the higher opportunity costs of unintended childbearing, so education affected negatively by affecting the preference for having children.

The study by Wang and Famoye (1997) was based on the "Chicago-Columbia" model. They used US data from the Michigan Plan Study of Income Dynamics for 1968 and for 1989. They use number of children in a family as dependent variable, which is over-dispersed.

They included women employment status, women education, family income, ethnicity (white or non-white) dummy, and rural-urban dummy as independent variables. They estimated both Poisson and negative binomial models and found that the standard errors of estimates are higher in Poisson model than generalized Poisson model, although both models give similar value for estimates. They find that mother's education, employment status, and family income have negative effects on fertility. They also find that fertility of non-whites is higher than that of whites.

Atella and Rosati (2000) investigated whether fertility depends on expected survival rate of children and the uncertainty associated with it. They analyzed data from the 1994 Human Development of India Survey. They used number of child births of women as the dependent variable. They included women and their husband age, marriage duration, wealth, education of women and husband, religion, male and female wage rate, number of children death, expected survival rate of children, and uncertainty of child survival rate.

They used village mean survival rate of children age under 5 as expected survival rate of children, and village variance of survival rate of children age under 5 as uncertainty of child survival rate of children. They used both Poisson and Poisson hurdle models.

Miranda (2010) argues that the double hurdle model is more appropriate for fertility data of Mexico than the single hurdle model. He mentioned that socio-economic characteristics affect women's fertility decision to transition from low to higher birth order in Mexico. He employed the double hurdle Poisson model using data from the 1997 Mexican Survey of Demographic Dynamics and used the number of children ever born to a woman who completed fertility as the dependent. He included women's age, women birth location, religion (Catholic), women's education, and ethnicity, as explanatory variables.

For the model, he constructs two hurdles, first at number zero and second at number three. He founds that education and Catholicism reduces the probability for women to have more than three children. For women having more than three children born in the south region, he found that Catholics have higher fertility than non-Catholic, and indigenous language speaker have fertility than non-indigenous language speakers.

Kravdal (2002) estimates a discrete-time Hazard regression model using the Demographic and Health Surveys for twenty two Sub-Saharan countries. He carried out Monte Carlo simulation to examine how educational distribution affects total fertility of women. He estimates two separate models for twenty two countries. The first is for women having first birth and the second for women having more than one birth. In the first model, he follows women who do not have child until first birth in three months interval for two years period. In the same way, he follows the women who have higher order birth.

He uses individual women education, average length of education of each community, rural versus urban, proportion of Muslim, proportion of other religion, and wealth indicator as explanatory variables. He found that the women's education strongly affects the first order birth, but the effect becomes weaker in higher order birth. He also obtains average length of education of the community influences fertility of women negatively.

Sennott and Yeatman (2012) in their study conducted in Malawi discovered that events that change one's economic circumstances might alter plans for future childbearing. For instance, job loss could lead to postponement of pregnancy to allow time for a household to regain financial balance before adding another member. On the contrary, a spouse beginning a new job could hasten a woman's childbearing plans. Frequent changes in fertility preferences may also reflect the economic uncertainty that is common in developing societies (Johnson-Hanks 2005, 2007; Agadjanian, 2005) such as Malawi, where employment may be sporadic or scarce. A significant relationship between occupation and desired fertility and fertility-related behavior is evident in several studies. Urban people prefer smaller families. Family size preference also varies regionally with variations of place of residence (Ali, 2000). Regional variation exists in regard to fertility intention because of different sociocultural pattern and practices. An analysis of survey data from 17 Arab states suggested that the fertility transition in most countries is being led by urban and literate women (Farid, 1996). Sidze et al (ND) in their study carried out among women in both rural and urban Senegal arrived at the finding that; age at first marriage occurs early in Senegal. In urban areas, over 49 percent of women aged 40-49 years were married before age 20 and 53 percent among urban women aged 15-29 years. On the other hand, 71 percent of rural women aged 40-49 years and 82 percent of rural women aged 15-29 were married before age 20. Early ages at first marriage expose Senegalese women to a long duration of pregnancy risk and high odds to give birth to numerous children.

Ayehu (1998) in his study among the Meru of Kenya discovered that women married to husbands with higher occupation status were more likely to desire to stop childbearing than those married to husbands with lower or middle status occupation hence an inverse relationship between the desire for more children and occupation.

In sum, the literature seems to agree that the following factors are important fertility determinants: wages of women, household income, parental education, biological supply factors (example; starting age of cohabitation), social factors (example, child sex preference, community/ethnicity effect, preference on certain number of children, rural versus urban).

3. Methodology

3.1. Source of the Data

The source of the data for this study was the 2011 Ethiopia Demographic and Health Survey (EDHS). The 2011 Ethiopia Demographic and Health survey was implemented by the Central Statistical Agency (CSA) with a support from the ministry of Health. This was the third Demographic and Health Survey (DHS) conducted in Ethiopia, under the worldwide measure DHS project, a USAID-funded project providing support and technical assistance in the implementation of population and health surveys in countries worldwide.

3.1.1. Sample implementation in EDHS

The sampling frame used for EDHS 2011 is the Population and Housing Census (PHC) conducted in 2007 by the Central Statistical Agency (CSA, 2008). CSA has an electronic file consisting of 81,654 Enumeration Areas (EA) created for the 2007 census for 10 of its 11 geographic regions. Ethiopia is divided into 11 geographical regions. Each region is sub-divided into zones, each zone into Woredas, each Woreda into towns, and each town into Kebeles. Among the 81,654 EAs, 21 percent are in urban areas and the rest 79 percent are in rural areas.

The sample design for EDHS 2011 used a two stage stratified cluster sampling design. Stratification was achieved by separating each region into urban and rural areas. In total, 23 sampling strata were created because Addis Ababa region is entirely urban. The sample points were selected independently in each sampling stratum, by a two-stage selection. Among the first stage, 624 selected EAs, 187 were in urban areas and 437 were in rural areas and also the second stage 18,720 households, 5,610 are in urban areas and 13,110 are in rural areas. In the selected households, 15,908 eligible men were identified for individual interview and 16,663 completed interviews with women age 15-49, 5,514 in urban areas and 11,149 in rural areas. The analysis on desired number of children presented in this study was based on the 14,751 women aged 15-49 years included in the survey.

3.2. Variables in the Study

The response variable of this study, Y_i, is a count variable, the total number of desired children a woman at the reproductive age (15-49) would like to have in her lifetime, Y_i= 0, 1, 2, … where refers to the i^{th}individual mother.

Women age group has been used as a valuable factor for assessing fertility preference or desired number of children among women.

Factors included as potential predictors of ideal number of children by women at reproductive age (15-49) are presented in Table 1

Predictor Variables	*Representation of variable*	*Categories*
Place of residence	X_1	$0 = Urban$
		$1 = Rural$
contraception method	X_2	$0 = Never\ used$
		$1 = Ever\ used$
women of education	X_3	$0 = Illiterate$
		$1 = Primary$
		$2 = Secondary$
		$3 = Higher$
Women occupation	X_4	$0 = Not\ working$
		$1 = Working$
Wealth index	X_5	$1 = Poorest$
		$2 = Poorer$
		$3 = Middle$
		$4 = Richer$
		$5 = Richest$
Religion	X_6	$0 = Orthodox$
		$1 = Catholic$
		$2 = muslim$
		$3 = protestant$
		$4 = traditional$
		$5 = others$
Women's age	X_7	$1 = 15 - 20$
		$2 = 20 - 25$
		$3 = 25 - 30$
		$4 = 30 - 35$
		$5 = 35 - 40$
		$6 = 40 - 45$
		$7 = 45 - 49$
women living children	X_8	$0 = has\ living\ children$
		$1 = has\ not\ living\ children$

3.3. Methods of Data Analysis

3.3.1. Statistical Weight

3.3.1.1. Introduction to Design Weighting

Design weighting is a mess. It is not always clear how to use weights in estimating anything more complicated than a simple mean or ratios, and standard errors are tricky even with simple weighted means. Recently, the use of statistical weights has become increasingly prominent in statistics to adjust the distribution of the remaining subjects ‘characteristics to that of the target population (e. g. Rotnitzky and Robins 1997; Preisser et al. 2000;Yung and Rao 2000; Miller et al. 2001; Smith 2001). In order for any statistical inferences drawn from the survey data to be valid, this representativeness of the sample must be taken into account. In general terms, sampling weights are used to make the sample more representatives the target population. All analyses should use the sampling weights calculated for each interviewed household and for each interviewed individual. A sampling weight is an inflation factor which extrapolates the sample to the target population. Generally, weights are important (Leslie_Kish-Survey_Sampling,1995): Weighting is a necessary process that guarantees unbiased estimates for population parameters in complex study settings, equally representative for every individual within a sampling stratum, Weighting brings the sample population to the same scale of the target population with design weight, weighting keeps the sample distribution close to the distribution of the target population, especially when oversampling is applied in certain areas, reduced sampling errors and Weighting for correcting nonresponse.

3.3.1.2. Sampling Weight Methodology of women in DHS

Due to the power allocation of the sample to the different regions and to their urban and rural areas, sampling weights are required for any analysis using complex survey design like 2011 EDHS data to ensure representativeness of the survey results at all Reporting level. Since the 2011 EDHS sample is a two-stage stratified cluster sample, sampling weights were calculated based on sampling probabilities separately for each sampling stage and for each cluster. We use the following notations**:**

P_{1hi}: The selection probability of PSU i in stratum h

P_{2hij}: The selection probability of a HH j of PSU i in stratum h

Let a_h be the number of clusters selected in stratum h, M_{hi} the number of households according to the sampling frames in the i^{th} cluster, and $\sum M_{hi}$ hi the total number of households in the stratum.

The probability of selecting the i^{th} cluster in the 2011 EDHS sample was calculated as follows: $\frac{a_h M_{hi}}{\sum M_{hi}}$

Let b_{hi} be the proportion of households in the selected segment compared to the total number of households in the EA i in stratum h if the EA is segmented, otherwise b_{hi}=1. Then the probability of selecting cluster i in the sample is:

$$P_{1hi} = \frac{a_h M_{hi}}{\sum M_{hi}} \text{ x} b_{hi}$$

Let L_{hi} be the number of households listed in the household listing operation in cluster i in stratum h, let g_{hi} be the number of households selected in the cluster. The second stage's selection probability for each household in the cluster is calculated as follows:

$$P_{2hij} = \frac{g_{hi}}{L_{hi}}$$

The overall selection probability of each household in cluster i of stratum h is therefore the production of the two stages selection probabilities:

$$P_{hi} = P_{1hi} \text{ x } P_{2hij}$$

The design weight for each household in cluster i of stratum h is the inverse of its overall selection probability: $d_{hij} = \mathrm{W_{hij}} = \frac{1}{\mathrm{P_{hi}}}$

The differences of the household sampling weights and the individual sampling weights are introduced by individual non-response. The final sampling weights (both household and individual weights) were normalized in order to give the total number of unweight cases equal to the total number of weighted cases at the all reporting levels. The normalized weights are relative weights which are valid for estimating means, proportions and ratios, but not valid for estimating population totals and for pooled data. Let $\mathrm{P_{3hijk}}$ be the selection probability of a member from all the eligible members in HH j of PSU i of stratum h. The design weight for individual k (women or men) from household j is given by

$$\mathrm{d_{hijk}} = \frac{1}{\mathrm{P_{1hi}P_{2hij}P_{3hijk}}}$$

If only one individual is selected per HH, the selection probability is

$$\mathrm{P_{3hijk}} = \frac{1}{\text{(\#of eligible members in HH j)}}$$

The design weights are the inverse of the element's overall selection probabilities the chance the unit was selected in the sample. If we have a perfect frame (no frame deficiencies) and full response (no nonresponse), the design weight is also the final survey weight otherwise, we used adjustment by nonresponse. Individual sampling weights are calculated based on HH sampling weight with correction for individual or women nonresponse: $W_{Ihi} = \frac{W_{Hhi}}{R_{Ih}} = d_{hij} * \frac{1}{R_{Hh}} * \frac{1}{R_{Ih}}$

Where W_{Hhi} the Adjusted Design Weight for HH nonresponse: $W_{Hhi} = \frac{d_{hij}}{R_h} = \frac{1}{P_{1hi}P_{2hij}R_h}$

$\mathrm{R_h}$ is The HH response rate in stratum h:

$$\mathrm{R_h} = \frac{\sum_{\mathrm{i}} \mathrm{n_{hi}}^* \mathrm{d_{hij}}}{\sum_{\mathrm{i}} \mathrm{n_{hi}}\, \mathrm{d_{hij}}}$$

$\mathrm{n_{hi}}$ and $\mathrm{n_{hi}}^*$ are the number of HHs found and completed in cluster i of stratum h, respectively

where $\mathrm{R_{Ih}}$ is the individual response rate in stratum h

$$R_{Ih} = \frac{\sum_i I_{hi}^* W_{Hhi}}{\sum_i I_{hi}\, W_{Hhi}}$$

I_{hi} *and* I_{hi}^* are the number of individuals found and completed in cluster i of stratum h, respectively.

3.3.2. Statistical Models

In this study, the variable of interest is a count variable. When the response or dependent variable (Ideal number of children by a woman in this study) is a count (which can take on non-negative integer values (0, 1, 2, ...), it is appropriate to use non-linear models based on non-normal distribution to describe the relationship between the dependent variable and a set of predictor variables. For count data, the standard framework for explaining the relationship between the outcome variable and a set of explanatory variables includes the Poisson and Negative Binomial regression models. Unlike linear regression, count data regression models have counts as the response variable that can take only nonnegative integer values.

Numerous models have been developed specifically for count data (Long &Freese, 2006; Sano, Jeong, Acock, & Zvonkovic, 2005). These models can handle non-normality on the dependent variable and do not require the researcher to either dichotomize or transform the dependent variable. We shall focus on four of these models (Atkins & Gallop, 2007; Long &Freese, 2006; Sano et al., 2005): Poisson, Negative Binomial, Zero-inflated Poisson (ZIP), and Zero-inflated Negative Binomial (ZINB).

3.3.2.1. Poisson Regression Model

A Poisson regression model allows modeling the relationship between a Poisson distributed response variable and one or more explanatory variables. It is suitable for modeling the number of events that occur in a given time period or area. The Poisson distribution becomes increasingly positively skewed as the mean of the dependent variable decreases (Long &Freese, 2006), reflecting a common property of count data.

According to Sturman (1999), the apparent simplicity of Poisson comes with two restrictive assumptions. First, the variance and mean of the count variable are assumed to be equal. The other restrictive assumption of Poisson models is that occurrences of the specified event/behavior are assumed to be independent of each other. Poisson Regression Model provides a standard framework for the analysis of count data. Let Y_i represent counts of events occurring in a given time or exposure periods with rate μ_i Y_i are Poisson random variables with p.m.f.

$$P(Y_i = y_i, \mu) = \frac{e^{-\mu_i}\,{\mu_i}^{y_i}}{y_{i\,!}}$$

μ_i>0, i=1, 2, .n

$$y_i = 0, 1, 2, \ldots$$

where, y_i denotes the ideal number of children for the i^{th} woman in a given time orexposure periods with mean parameter μ_i

In Poisson model, the conditional variance is equal to conditional mean:

E (Y_i) =Var (Y_i) = μ_i

This property of the Poisson distribution is known as equidispersion.

Let x be a $n \times (p+1)$ covariate matrix**.** The relationship between $\boldsymbol{Y_i}$ and i^{th} row vector of x, x_i is given by the Poisson log-linear model

$$\ell\text{n}(\mu_i) = {x_i}^T\beta = \eta_i$$

where, $x_i = (1, x_{i1}, x_{i2}, \dots\dots\dots x_{ip})^T$ is the vector of explanatory variables and $\beta = (\beta_0, \beta_1, \dots, \beta_p)^T$ is the vector of the unknown regression parameters.

The regression parameters are estimated using the maximum likelihood estimation. The likelihood function of the Poisson model based on a sample of n independent observations is given by

$$\ell(\beta, Y_i) = \prod_{i=1}^{n} \frac{e^{-\mu_i} \mu_i^{y_i}}{y_{i\,!}}$$

The log-likelihood function is

$$l = \log(L(\beta)) = \sum_{i=1}^{n} [y_i \ln \mu_i - \mu_i - \ln y_i!]$$

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. Thus, we obtain the first derivatives of ℓ with respect to the underlying parameters as follows:

$$\frac{\partial l(\beta)}{\partial \beta_i} = \sum_{i=1}^{n} (y_i - \mu_i) x_{ij}$$

If $E(\boldsymbol{y_i}) < var(\boldsymbol{y_i})$, then we speak about over- dispersion, and when $E(\boldsymbol{y_i}) > var(\boldsymbol{y_i})$, we say we have under-dispersion. Next, we employed two tests of over dispersion where the Null Hypothesis(H_0) is: mean and variance of the response variable are equal against the Alternative Hypothesis(H_1): variance exceeds the mean. There are two basic criteria commonly used to check the presence of over dispersion:

1. Deviance, $D(y, \hat{\mu})$, is given by

$$D(y, \hat{\mu}) = 2 * \sum_{i}^{n} \{y_i \ln\left(\frac{y_i}{\widehat{\mu_i}}\right) - (y_i - \widehat{\mu_i})\}$$

Where, y is the number of events, n is the number of observations and $\widehat{\mu_i}$ is the fitted Poisson mean.

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

$$x^2 = \sum_{i=1}^{n} \left(\frac{(y_i - \widehat{\mu_i})^2}{\widehat{\mu_i}} \right)$$

Over-dispersion may be a result of higher occurrence of zero counts and subject heterogeneity.

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 variance is an over dispersion, while values smaller than one indicate an under-dispersion. It is possible to account for over-dispersion with respect to the Poisson model by introducing a scale (dispersion) parameter into the relationship between the variance and the mean (Pedan, 2001). Another way of checking the presence of over-dispersion is a statistical test of the hypothesis:

$$H_0: \alpha = 0 \;\; Vs \; H_1: \alpha > 0.$$

If P-value of LRTα < (level of significance), then there is over-dispersion and the Negative Binomial model is preferred.

The Negative Binomial Regression Model is more appropriate for over-dispersed data because it relaxes the constraints of equal mean and variance.

In the general, Poisson Regression Model, we think of μ_i as the expected desired number of children from the i^{th}mother woman and the total number live birth children from the i^{th}mother woman is Ni. This means, parameter will depend on the population size and the total number of live birth children from the individual mother. Thus, the distribution of Yi can be written as:

$$Y_I \sim Poisson(N_I \mu_I)$$

Where N_I are the total fertility rate of i^{th} mother and $\mu_I = exp^{(X_i{}^T \beta)}$the logarithm of the children birth live is introduced in the regression model as an offset variable. By including

$$\log \mu_i = \log N_i + X_i{}^T \beta$$

The link between the expectation of the dependent variable and the linear predicator is a logarithmic function and the linear predictor contains a known part or offset. This allows for estimation of maximum likelihood, standard errors and the likelihood ratio goodness of fit chi-square statistics (Agresti, A. 2008). The model suggests that both set of the parameters are dependent on the covariates.

Furthermore, the number of children born will be equal to the observed deaths if the coefficients of the independent variables, denoted by β, are all equal to zero. Since $\log N_i$ is a constant, any variation in the coefficients of the independent variables will show up affecting the dependent variable and not the number of children born. The procedure therefore allows us to obtain the maximum likelihood regression coefficients that can be easily interpreted in terms of differentials in the dependent variables. Using the Negative Binomial Regression procedure, several regression equations are estimated to the relationship between under-five mortality changes when control variables earlier mentioned are introduced. Results from the Negative Binomial Models are sometimes better expressed on more convenient scale.

3.3.2.2. Negative Binomial Regression Model

The NB Regression Model is used when count data are over dispersed (i.e when the variance exceeds the mean). Over dispersion, caused by heterogeneity or an excess number of zeros (or both) to some degree is inherent to most Poisson data. By introducing a random component into the conditional mean, the Negative Binomial Regression Model addresses the issue of over-dispersion. However, it equally models both zero and nonzero counts, which might result in a poor fit for data with excessive number of zeros. Therefore, it is always necessary to check the proportion of zero counts before developing a Negative Binomial Regression Model.

This study used the likelihood ratio test to determine the more appropriate model between the Poisson Regression and Negative Binomial Regression. Hilbe

(2007) used Negative Binomial Regression to Model over dispersed Poisson data. When the Negative Binomial is used to model over-dispersed Poisson count data, the distribution can be thought of as an extension to the Poisson Model. The Negative Binomial Regression Model uses a log link function between the dependent variable(Ideal number of children of woman) and independent variables.

The only difference between the Poisson and the NB lies in their variances, regression coefficients tend to be similar across the two models, but standard errors can be very different. The NB regression model is

$$P(y_{i\,,}\, \mu_{i\,,\alpha}) = \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_{i\,!}\Gamma\left(\frac{1}{\alpha}\right)}(1 + \alpha\mu_i\,)^{\frac{-1}{\alpha}}\left(1 + \frac{1}{\alpha\mu_i}\right)^{-y_i}$$

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

With mean and variance are given by

$E(Y_i) = \mu_i = exp^{(X_i{}^T\beta)}$ $and\ Var(Y_i) = \mu_i(1 + \alpha\mu_i)$

Where, α shows the level of over-dispersion and Γ (.) is the gamma function. If$\alpha = 0$, NB Regression Model will reduce to Poisson Regression Model. Often data will show over-dispersion (Variance > mean) or under-dispersion (Variance < mean). With over-dispersed data we may well use the Negative Binomial Regression model.

This Model adds unobserved heterogeneity by specifying

$\mu_i = E(y_i) = exp^{(X_i{}^T\beta)}$

Where, $X_i{}^T$ is $1xp$ row vector of covariate (including an intercepts), p is the number of covariate in the model and $px1$ column vector of unknown regression parameters.

The likelihood function of the NB model based on a sample of n independent observations is given by

$$L(\mu,\alpha,y_i)=\prod_{i=1}^{n}\{\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i\,!\ \Gamma\left(\frac{1}{\alpha}\right)}(1+\alpha\mu_i)^{\frac{-1}{\alpha}}\left(1+\frac{1}{\alpha\mu_i}\right)^{-y_i}\}$$

The log-likelihood function ℓ of NB regression model is

$$\ell=\sum_{i=1}^{n}\left\{-\log(y_i!)+\sum_{k=1}^{y_i}(\alpha y_i-\alpha k+1)-\left(y_i+{}^{1}/_{\alpha}\right)\log(1+\alpha\mu_i)-y_i\log(\mu_i)\right\}$$

Where, $\frac{\Gamma\left(y_i+\frac{1}{\alpha}\right)}{y_i\,!\,\Gamma\left(\frac{1}{\alpha}\right)}=\prod_{k=1}^{y_i}\left(y_i+\frac{1}{\alpha}-k\right)=\alpha^{-y_i}\prod_{k=1}^{y_i}(\alpha y_i-\alpha k+1)$

For estimating regression coefficients β and dispersion parameter α, the Newton-Raphson iteration procedure is applied like in the Poisson model.

3.3.2.3. Zero-inflated Regression Model

Real-life count data are frequently characterized by over-dispersion and excess zeros. Zero inflated count models provide a parsimonious yet powerful way to model this type of situation. Such models assume that the data are a mixture of two separate data generation processes: one generates only zeros, and the other is either a Poisson or negative binomial data-generating process Count data that have an incidence of zeros greater than expected for the underlying probability distribution can be modeled with a zero inflated distribution. The population is considered to consist of two subpopulations. Observations drawn from the first subpopulation are realizations of a random variable that typically has either a Poisson or Negative Binomial distribution, which might contain zeroes.

3.3.2.3.1. Zero- inflated Poisson Regression Model

Suppose the mean of the underlying Poisson distribution is μ and the probability of an observation being drawn from the constant distribution that always generates zeros is ω_i. The parameter ω_i is often called the zero inflation probability. (Agresti, A., 2002)

The probability distribution of a zero inflated Poisson random variable Y is given by

$$p(y_i) = \begin{cases} \omega_i + (1-\omega_i)e^{\mu_i}\,, y_i = 0 \\ (1-\omega_i)\dfrac{e^{-\mu_i}\ \mu_i{}^{y_i}}{y_i\,!} & , y_i = 1,2,3\ldots \end{cases} \quad 0 \le \omega_i \le 1$$

Where, $Y_i \sim ZIP(\mu_i, \omega_i)$. the mean and variance of ZIP are given by

$$E(Y_i) = (1-\omega_i)\mu_i \ and\ Var(Y_i) = E(Y_i)\big((1-\omega_i)\mu_i\ \big)$$

The excess zeros are a form of over-dispersion and fitting a zero inflated Poisson model can account for the excess zeros, but there are also other sources of over-dispersion that must be considered. If there are sources of over-dispersion that cannot be attributed to the excess zeros, failure to account for them constitutes a model misspecification, which results in biased standard errors. In ZIP Models, the underlying Poisson distribution for the first subpopulation is assumed to have a variance that is equal to the distribution's mean. If this is an invalid assumption, the data exhibit over-dispersion (or under-dispersion). (Agresti, A., 2002) A useful diagnostic tool that can aid in detecting over dispersion is the Pearson chi-square statistic defined as

$$x^2 = \sum_i \frac{(y_i - \mu_i)^2}{V(\mu_i)}$$

Comparing the computed Pearson chi-square statistic to an appropriate quintile of a chi-squared distribution with n-p df constitutes a test of over-dispersion. If over dispersion is detected, the ZINB Model often provides an adequate alternative.

3.3.2.3.2. Zero-inflated Negative Binomial Regression Model

Zero-inflated Negative Binomial Regression is for modeling count variables with excessive zeros and it is usually for over-dispersed count outcome variables. Furthermore, theory suggests that the excess zeros are generated by a separate process from the count values and that the excess zeros can be modeled independently. The probability distribution of a Zero inflated Negative Binomial random variable Y is given by

$$p\left({}^{y_i}/_{w}\ ,\alpha,\mu\right)=\begin{cases}\omega_i+(1-\omega_i)\left(1+\alpha\mu_i\right)^{-\frac{1}{\alpha}}, & y_i=0\\ (1-\omega_i)\dfrac{\Gamma\left(y_i+\frac{1}{\alpha}\right)(1+\alpha\mu_i)^{\frac{-1}{\alpha}}\left(1+\frac{1}{\alpha\mu_i}\right)^{-y_i}}{y_i\,!\Gamma\left(\frac{1}{\alpha}\right)}, & y_i>0\end{cases}$$

Where, μi is the mean of the underlying negative binomial distribution, α >0 is the over dispersion parameter and is assumed not to depend on covariates and $0 \leq \omega_i \leq 1$.Also the parameters μ_i *and* ω_i depend on vectors of covariates X_i *and* Z_i, respectively. (Agresti, A., 2002)

The formulations for μ_i *and* $\omega_i$are the same as those used in the zero-inflated Poisson regression model. In this case, the mean and variance of the *Y*i are

$$E(y_i)=(1-\omega_i)\mu_i$$
$$var(y_i)=(1-\omega_i)\mu_i(1+\omega_i\mu_i+\alpha\mu_i)$$

3.3.2.4. Parameter Estimation of ZINB Model

The ZINB distribution is not a standard GLM type exponential family distribution, even when the over dispersion parameter is known, and standard GLM fitting methods are not applied. To obtain the parameter estimates of ZINB regression models $\widehat{\alpha}, \widehat{\beta},\ \hat{\gamma}$, the Newton-Raphson method can be used.

The log likelihood function $l = l(\alpha, \mu_i, w_{i\,;}\, y)$, for the ZINB Model is given below.

$$\ell = \ell(\alpha, \mu_i, w_{i\,;}\, y)$$

$$= \sum_{i=0}^{n} \left\{ I_{(Y_i=0)} \log\left(w_i + (1-w_i)(1+\alpha\mu_i)^{\frac{-1}{\alpha}}\right) \right.$$

$$\left. + I_{(Y_i>0)} \log[(1-w_i) \frac{\Gamma\left(Y_i + \frac{1}{\alpha}\right)(1+\alpha\mu_i)^{\frac{-1}{\alpha}}\left(1+\frac{1}{\alpha\mu_i}\right)^{-y_i}]}{Y_i!\,\Gamma\left(\frac{1}{\alpha}\right)} \right\}$$

Since, $\frac{\Gamma\left(Y_i+\frac{1}{\alpha}\right)}{Y_i!\Gamma\left(\frac{1}{\alpha}\right)} = \prod_{k=1}^{Y_i}\left(Y_i + \frac{1}{\alpha} - k\right) = \alpha^{-Y_i}\prod_{k=1}^{Y_i}(\alpha Y_i - \alpha k + 1)$

Furthermore, it can be written as

$$\ell = \sum_{i=0}^{n} \left\{ I_{(Y_i=0)} \log\left[w_i + (1-w_i)(1+\alpha\mu_i)^{\frac{-1}{\alpha}}\right] + I_{(Y_i>0)} \log[(1-w_i) - \log y_i! \right.$$

$$\left. + \sum_{k=1}^{y_i}(\alpha Y_i - \alpha k + 1) - \left(Y_i + \frac{1}{\alpha}\right)\log(1+\alpha\mu_i) + \log y_i! + y_i \log\mu_i] \right\}$$

Newton-Raphson iteration procedure can be used for estimating the parameter of ZINB Regression Models.

3.3.2.5. Goodness of fit tests

3.3.2.5.1. Likelihood Ratio test

The Likelihood ratio test 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:

Likelihood ratio test $G^2 = -2(l_{null} - l_k) \sim X_{P-1}^2$

This statistic is called the likelihood-ratio test statistic.

Where: l_{null} is the log-likelihood of the null model and

l_k is the log-likelihood of the full model

Comprising k predictor, p is number of parameters and X_{P-1}^2 is a chi-square distribution with p-1 degree of freedom. If the test statistics exceeds the critical value, the null hypothesis is rejected. That means the overall model is significant.

In this study, to compare Poisson and NB regression models and also ZIP with ZINB regression models, we used significance of dispersion parameter and likelihood ratio (LR) test as criterions. The statistic of likelihood ratio test for α is given by the following equation: $LRT_\alpha = -2(LL_1 - LL_2)$

This statistic has a Chi-squared distribution with 1 degrees of freedom and LL is log-likelihood. If the statistic is greater than the critical value then, the model 2 is better than the model 1.

3.3.2.5.2. Vuong Test

The Vuong test is a non-nested test that is based on a comparison of the predicted probabilities of two models that do not nest (Vuong, 1989). That means vuong test statistics are needed to provide the appropriateness of zero-inflated models against the standard count models. For instance, comparisons between zero-inflated count models with ordinary Poisson, or zero-inflated negative binomial against ordinary negative binomial model can be done using Vuong test. This test is used for model comparison.

For testing the relevance of using zero-inflated models versus Poisson and NB regression models, the Vuong statistic is used. Let's define

$$m_i = log\left(\frac{P_1\left(\frac{y_i}{x_i}\right)}{P_2\left(\frac{y_i}{x_i}\right)}\right)$$

Where, $P_1\left(\frac{y_i}{x_i}\right)$and $P_2\left(\frac{y_i}{x_i}\right)$ are probability mass functions of zero-inflated and Poisson or NB models, respectively. In general, $P_N\left(\frac{y_i}{x_i}\right)$ is the predicted probability of observed count for case i from model N, then the Vuong test statistic is simply the average log-likelihood ratio suitably normalized.

The test statistic is

$$V = \sqrt{n}\frac{\frac{\sum_{i=1}^{n} m_i}{n}}{\sqrt{\frac{\sum_{i=1}^{n}(m_i - \overline{m})^2}{n}}} = \frac{\sqrt{n}}{s_m}(\overline{m})$$

Where, $\overline{m}$ are mean of m_i, s_m standard deviation and n sample size

The hypotheses of the Vuong test are:

$H_0: E[m_i] = 0 \quad H_1: E[m_i] \neq 0$

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

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

3.3.2.5.3. AIC and BIC

AIC and BIC are goodness of fit criteria used for model selection. The likelihood ratio test was used to compare the Poisson model and NB model. Many Monte-Carlo simulations indicate that the BIC and AIC selection criteria need to be used together [Dalrymple et al (2003) and Wang et al (1996)]. The model with smallest value of AIC or of BIC is preferable. Selecting an appropriate model is often based on a standard likelihood information criteria, for example, Akaike information criteria (Akaike, 1973) or Bayesians information criteria (Raftery, 1986) abbreviated by AIC and BIC, respectively, Where

AIC = -2 log likelihood + 2k

BIC = -2 log likelihood +k ln (n)

Where, k = number of parameters and n = number of observations.

3.3.2.5.4. Test for individual predictors

Let β denote an arbitrary parameter. Consider a significance test of H_0: β = 0.The simplest test statistic uses the large-sample normality of the ML estimator β, let SE(β) denote the standard error of β, evaluated by substituting the ML estimate for the unknown parameter in the expression for the true standard error. When H_0 is true, the test statistics

$$Z = \frac{\hat{\beta}-\beta_0}{SE(\hat{\beta})}$$

Has approximately a standard normal distribution. Equivalently, z^2 has approximately a chi-squared distribution with df = 1. This type of statistic, which uses the standard error evaluated at the ML estimate, is called a Wald statistic. The Wald statistic is $Z^2 = (\frac{\hat{\beta}-\beta_0}{SE(\hat{\beta})})^2$

Under H_0 true, Z^2 is a chi-square distribution with 1 degree of freedom. Wald statistics are for small samples. Likelihood-ratio tests are generally considered to be superior (Agresti, 2007).

4. Results and Discussion

This chapter discusses results of the study showing how selected socio economic, and demographic factors/variables affect desired number of children among women at reproductive age in Ethiopia.

The statistical analyses were performed using South Texas Art Therapy Association (STATA) version 13/14, Statistical Package for Social Science (SPSS) version 23, and Microsoft-Excel.

4.1 Descriptive Statistics

In order to have an overall picture of the distribution of the Ideal number of children over lifetime by women at reproductive age (15-49), all descriptive analysis were performed using weighted data.

Table 4.1 shows the descriptive statistics of the weighted data corresponding to the dependent variable and all explanatory variables. The variance of the outcome variable (desired number of children), 7.40, is greater than the population mean, 4.30 and the ratio $\frac{7.40}{4.30} = 1.72 > 1$ the results suggesting over dispersion.

The Table also shows that the maximum number of ideal number of children reported by the women in the weighted data is 20 while the minimum is 0. Out of a total of 14751 women in the weighted data, 48 percent were uneducated, 40 percent, 7 percent and 5 percent had primary, secondary and higher education respectively.

The proportion (%) of women in the poorest (17.1 percent) and poorer (17.8 percent) wealth index are lower than that for middle (18.3 percent), richer (20.0 percent) or richest (26.9 percent) wealth index.

Table 4.1. Proportion of women in each category of variables

Variable	Obs	Proportion
women age in five year		
15-19	3609	0.260
20-24	2802	0.184
25-29	2891	0.194
30-34	1822	0.122
35-39	1635	0.108
40-44	1087	0.069
45-49	905	0.064
women education		
no_educ	6932	0.479
primary_edu	5497	0.398
secod_edu	1363	0.075
high_edu	959	0.049
women living in place		
urban	4972	0.252
rural	9779	0.749
women Religion		
Orthodox	6422	0.481
catholic	167	0.011
protestant	2729	0.225
muslim	5245	0.269
traditional	74	0.007
other	114	0.007
women wealth index		
poorest	3094	0.171
poorer	2082	0.178
middle	2018	0.183
richer	2284	0.200
richest	5273	0.269
women use of contracep		
not_use_co~a	11991	0.798
contra_use~s	2760	0.202
women currently working		
current_no~k	7022	0.417
currently_~g	7729	0.583
women has living child		
has_no_liv~d	5307	0.362
has_living~d	9444	0.638

Additionally, 25.2% of the women were living in urban areas and 22.4% were using contraception at the time of the survey.

Likewise, about 48.1 percent of the women in the weighted data set were orthodox Christians and about 26.9 percent were Muslim. The proportion of protestant, catholic, traditional and other religion followers was 22.4 percent, 1.1 percent, 0.7 percent and 0.7 percent respectively. The proportion of both traditional and others religion follower women were lower than that for catholic, protestant, Muslim and also orthodox women.

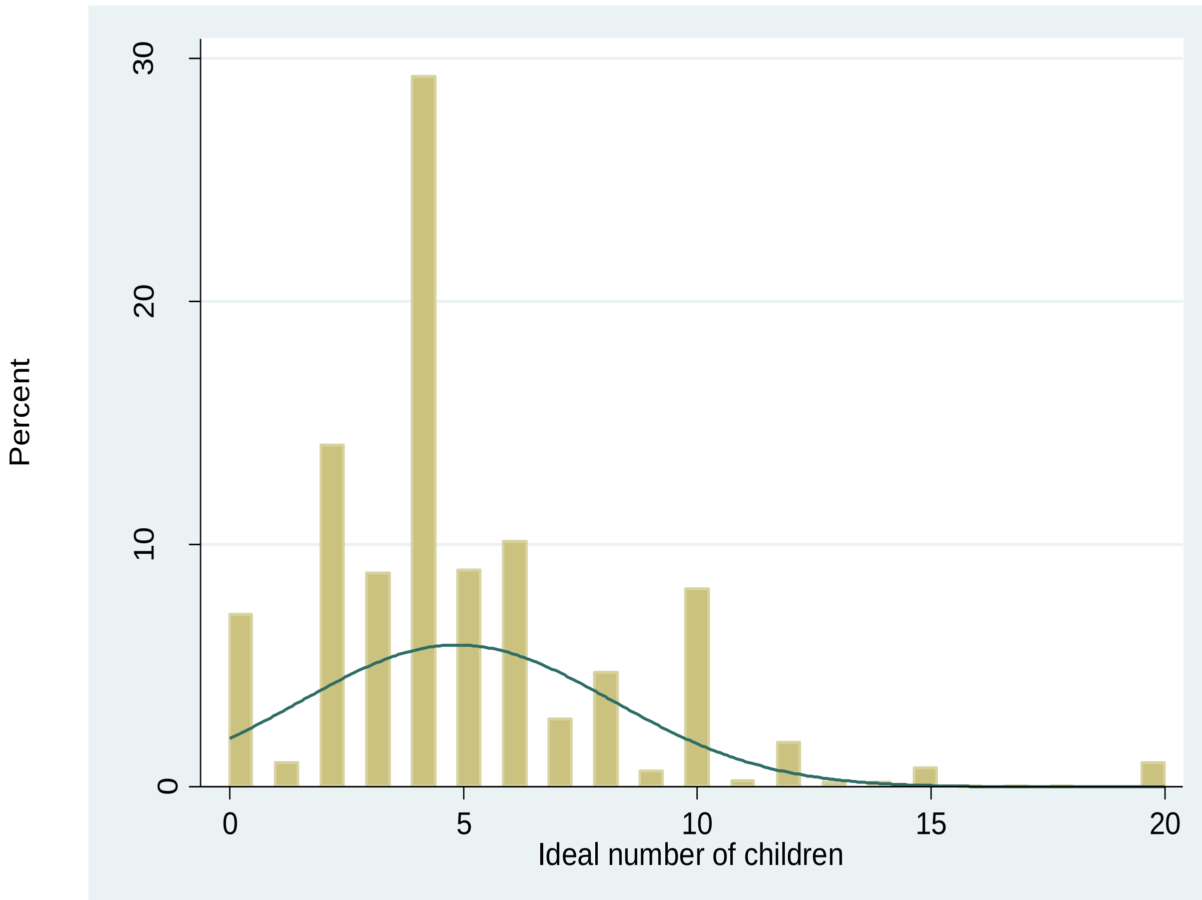
Table 4.2. Descriptive statistics of INC using weighted & unweighted Data

Variable	Category	Mean	Std.Dev.
Ideal number of children	Data is unweighted	4.86	3.33
	Data is weighted	4.3	2.72

Table 4.2 shows that mean and standard deviation of desired/ideal number of children by a woman at the reproductive age are lower for the weighted dataset than for the unweighted dataset. Using the weighted data will also reduce bias, standard error, sampling error. Descriptive analysis of data based on weights will provide more precise results than that without weights (Leslie_Kish-Survey_Sampling,1995).

Since, when the statistical results more precision becomes the measurements of standard error would be declined. Because, the sampling error is measured by standard error, coefficient of variation, margin of error, confidence interval.

As shown in Figure 4.1, the distribution of the desired number of children has a somewhat slowly decreasing tail and is skewed to the right.



histogram INC, percent normal
(bin=41, start=0, width=.48780488)

Figure 4.1. Histogram for desired number of children by women weighted

Summary Statistics of Selected Variables

Summary statistics for the Ideal number of children **(**INC**)** by women aged 15-49 years are presented in Table 4.3. The average desired number of children for the entire weighted data is 4.3. The data revealed substantial variation in this index when controls are introduced

Table 4.3.Mean and SD of desired number of children among Women 15-49

Variable	Mean of INC	Std. Dev.
Total desired children	4.302	2.723
women age in five year		
15-19	3.266	1.948
20-24	3.854	1.957
25-29	4.312	2.471
30-34	4.971	2.913
35-39	5.262	3.178
40-44	5.439	3.534
45-49	5.654	3.726
women education		
no_educ	5.006	3.184
primary_edu	3.754	2.123
secod_edu	3.374	1.598
high_edu	3.29	1.481
women living in place		
Urban	3.691	2.219
Rural	4.508	2.843
women Religion		
Orthodox	3.881	2.301
Catholic	4.282	2.597
Protestant	4.197	2.323
Muslim	5.109	3.439
Traditional	5.104	2.949
Other	4.839	3.146
women wealth index		
Poorest	4.919	3.183
Poorer	4.594	2.685
Middle	4.457	2.875
Richer	4.178	2.683
Richest	3.705	2.176
women use of contracep		

not_use_co~a	4.344	2.821
contra_use~s	4.137	2.285
women currently working		
current_no~k	4.417	2.948
currently_~g	4.22	2.547
women has living child		
has_no_liv~d	3.294	1.873
has_living~d	4.875	2.955

Women age: The average desired number of children was highest (5.65) among women in the oldest (45-49) age category and lowest (3.27) among women in the youngest age group (15-19 years)suggesting that average desired number of children is positively correlated with age. Or the older the respondent, the more children that she considers ideal; women age 15-19 respond that the ideal family size is 3.27 children, while women age 45-49 say it is 5.65.

Women education: - Clear difference in the ideal number of children has been observed between women with no education and those with education. Women with some education exhibited lower average desired number of children than those without education. That is, difference in the average Ideal number of children was observed between women who have never been to school and those with secondary and higher education. Average desired number of children of women with no education was 5. Conversely, the average desired number of children of women with primary, education, secondary education and higher education was 3.73, 3.37 and 3.29 respectively.

Place of residence: There was a difference in the average desired number of children by place of residence. The mean of desired number of children was higher in rural areas (4.5) than in urban areas (3.7). The ideal number of children among rural women is almost one child more than among urban women.

Working/employed women exhibited average desired number of children is 4.2 and also unemployed/ not working women 4.42. Women currently working, women who are employed exhibit lower fertility than those who are not.

Contraceptive use: The ideal number of children in contraceptive use 4.14 and women's not used contraceptive method 4.34 children's needed. Women use of contraceptive**;** they have the lowest mean value desired children of women who have ever used modern contraceptive methods than those who are not.

Religion:-The mean ideal number of children for Orthodox women (3.88) is lower than that for protestant (4.2), catholic (4.28), other religion (4.82) Muslim (5.1) and traditional (5.1) religion followers of women.

Household wealth index:

The ideal number of children was highest (4.9) for women in the poorest quintile and lowest (3.7) for those in the richest quintile suggesting that desired number of children is negatively correlated with wealth index. The ideal number of children wealth quintile increase. Women in the lowest wealth quintile consider (4.9) children to be ideal, compared with 3.7 children among women in the highest wealth quintile.

Women with living child: Table 4.3 shows that women without living children exhibited lower average ideal number of children (3.29) than the average desired number of children by women having living children 4.87.

Table 4.4: Desired number of children by place of residence and by women with/without living children

INC	Place of Residence		Women with/without living child		
	Urban	Rural	has no living	has living child	Total
0	180	1111	416	875	1291
1	54	102	98	58	157
2	921	1096	1442	576	2018
3	452	882	723	611	1334
4	1466	3551	1875	3141	5016
5	177	1133	360	950	1310
6	207	1367	263	1310	1574
7	49	329	36	342	378
8	71	588	50	609	659
9	5	73	3	76	78
10	64	562	56	570	627
11	0	20	1	19	20
12	40	111	13	138	151
13	2	16	0	18	18
14	2	7	2	7	9
15	10	51	3	59	61
16	0	1	0	1	1
17	0	2	0	2	2
18	1	1	1	1	1
20	7	40	5	42	47
Total	3709	11042	5346	9405	14751

Table 4.4 shows that 25 percent of the women were living in urban parts of the country and 75 percent of them were living in rural parts of the country. And also 36 percent of the women had no living children and 64 percent of them had living children during the survey period.

4.2. Count Regression Model Results

4.2.1. Parameter Estimation

Table 4.6 presents the fitted models (with estimated coefficients and their corresponding standard errors) for the four count regression models considered (Poison, Negative Binomial, Zero-inflated Poison, and Zero-inflated Negative binomial models). Also presented in Table 4.6 are over dispersion parameters alpha and its natural logarism. The results show that the over-dispersion parameter alpha is significantly different from zero indicating over-dispersion of the data. Moreover, the ratio of the Deviance and Pearson Chi-square statistic to their corresponding degrees of freedom are greater than one, indicating over-dispersion in the data and the Negative Binomial (NB) regression model is preferred over the Poisson model (Table 4.5).

Table 4.5.: The results of over-dispersion test after fitting a Poisson regression

Statistics	Value	Degrees of freedom	Value/Deg. Freedom	P-value
Deviance test statistics	22996.29	14728	1.5614	0.0000
Pearson Chi-square statistic	19366.53	14728	1.3149	0.0000

The fact that the Negative Binomial regression model is favored over the Poisson regression model has also been confirmed by the likelihood-ratio test. Since the value of this statistic is 2179.23 with $p < 0.0001$, we reject the null hypothesis that there is equi-dispersion, and conclude that there is significant over-dispersion in the data.

When the assumption of the equality of variance and mean in the Poisson regression model is violated, over-dispersion occurs and the standard error estimates will be biased which leads to incorrect value of the test statistic. Consequently, the covariates may be wrongly interpreted (Yaacob et. al., 2010).

Since both AIC and BIC values were lower for the Negative Binomial model than for the Poisson model, the Negative Binomial model is preferred. The Negative binomial model accounts for the over-dispersion in the data.

Table 4.6. Estimated Coefficients and Std.Err. Of the Poisson, NB, Zero-inflated and also alpha

Variable	Poisson		NB		ZIP		ZINB	
	Coef.	Robust Std. Err.	Coef.	Robust Std. Err.	Coef.	Robust Std.Err.	Coef.	Robust Std. Err.
women_age								
20-24	-0.138**	0.021	-0.135**	0.021	-0.163**	0.019	-0.163**	0.019
25-29	-0.299**	0.024	-0.298**	0.023	-0.298**	0.020	-0.298**	0.020
30-34	-0.331**	0.028	-0.333**	0.027	-0.341**	0.024	-0.341**	0.024
35-39	-0.431**	0.029	-0.431**	0.028	-0.403**	0.024	-0.403**	0.024
40-44	-0.553**	0.035	-0.550**	0.035	-0.495**	0.028	-0.495**	0.028
45-49	-0.633**	0.038	-0.630**	0.037	-0.553**	0.030	-0.553**	0.030
women_educ								
Primary	-0.066**	0.017	-0.067**	0.017	-0.103**	0.014	-0.103**	0.014
Secondary	-0.120**	0.025	-0.119**	0.025	-0.188**	0.023	-0.188**	0.023
Higher	-0.160**	0.030	-0.161**	0.029	-0.229**	0.028	-0.229**	0.028
Place_Resid								
Rural	0.013	0.031	0.017	0.029	0.055**	0.026	0.055**	0.026
Religion								
Catholic	0.053	0.057	0.055	0.056	0.066	0.049	0.066	0.049
Protestant	0.056**	0.016	0.062**	0.016	0.078**	0.013	0.078**	0.013
Muslim	0.223**	0.017	0.225**	0.017	0.250**	0.014	0.250**	0.014
Traditional	0.211**	0.089	0.218**	0.090	0.241**	0.069	0.241**	0.069
Other	0.093	0.082	0.103	0.082	0.195**	0.055	0.195**	0.055
wealth_index								
Poorer	-0.044**	0.022	-0.043**	0.022	-0.060**	0.018	-0.060**	0.018
Middle	-0.062**	0.024	-0.061**	0.023	-0.044**	0.019	-0.044**	0.019
Richer	-0.092**	0.024	-0.090**	0.024	-0.063**	0.019	-0.063**	0.019
Richest	-0.109**	0.035	-0.104**	0.034	-0.086**	0.029	-0.086**	0.029
contra_use								
yes	-0.093**	0.018	-0.092**	0.017	-0.106**	0.015	-0.106**	0.015
women_occup								
working	-0.052**	0.014	-0.047**	0.014	-0.056**	0.012	-0.056**	0.012
women with								
yes	0.134**	0.019	0.132**	0.019	0.135**	0.017	0.135**	0.017
_cons	-1.579**	0.040	-1.588**	0.038	-1.539**	0.033	-1.539**	0.033
ln(women_age)	1.000	(exposure)	1.000	(exposure)	1.000	(exposure)	1.000	(exposure)
/lnalpha			-2.383	0.086			-9.988	77.816
alpha			0.092	0.008			0.006	0.004

Notes: Robust Standard Error is in **Robust Std. Err.** ** Significant at 5% confidence
Alpha is dispersion parameter. $p - value < 0.05$

4.2.2. Comparison of Models

A critical question in data analysis is how to choose the appropriate models for a specific study. Several criteria can be used to compare and select among considered models. In this study, four different count regression models, namely; Poisson, negative binomial, zero-inflated Poisson and zero-inflated negative binomial models were considered. Different model selection criteria: the Log pseudo likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used in order to identify the most appropriate fitted model. In cases of over-dispersion, the ZIP model typically fits better than a standard Poisson model. But there are other models that allow for over dispersion: the standard negative binomial regression model and ZINB model. Since the Poisson and NB are not nested within the ZIP and ZINB models, respectively, we use the Vuong test to compare the non-nested models: ZIP versus Poisson and ZINB versus NB regression models.

Table 4.7. Model selection criteria for PR, NB, ZIP and ZINB models for the desired number of children by women weighted dataset

Selection criteria	Models			
	Poisson	**NB**	**ZIP**	**ZINB**
Log pseudo likelihood	-33768.068	-33274.915	-31507.04	-31507.04
AIC	67582.14	66597.83	63106.08	63108.08
BIC	67756.92	66780.21	63455.64	63465.25
Vuong			16.38	18.73
(p-value)			0.0000	0.0000

Table 4.7 shows the model selection criteria used to identify the best/preferred model among the four candidate models. First, the calculated value of the Vuong test statistic (16.38) for comparing ZIP versus Poisson model is greater than 1.96 implying that the ZIP model is preferred to the Poisson model for predicting the desired number of children by women.

Similarly, the calculated value of the Vuong test statistic for comparing ZINB versus NB models is 18.73, indicating that the ZINB model is preferred to NB regression model. Finally, to compare the ZIP and ZINB models, AIC, BIC and Log pseudo likelihood were used as shown in Table 4.7. The model with the smallest AIC, smallest BIC and largest Log pseudo likelihood is preferred. Since ZIP model has the smallest AIC, smallest BIC and maximum Log pseudo likelihood, ZIP model is the most appropriate and preferred model among the four models. Thus, the Zero-inflated Poisson regression model with the lowest value of AIC, lowest value of BIC and the highest value of Log pseudo likelihood is the most appropriate and preferred model for describing the ideal number of children that woman at reproductive age would like to have in their lifetime.

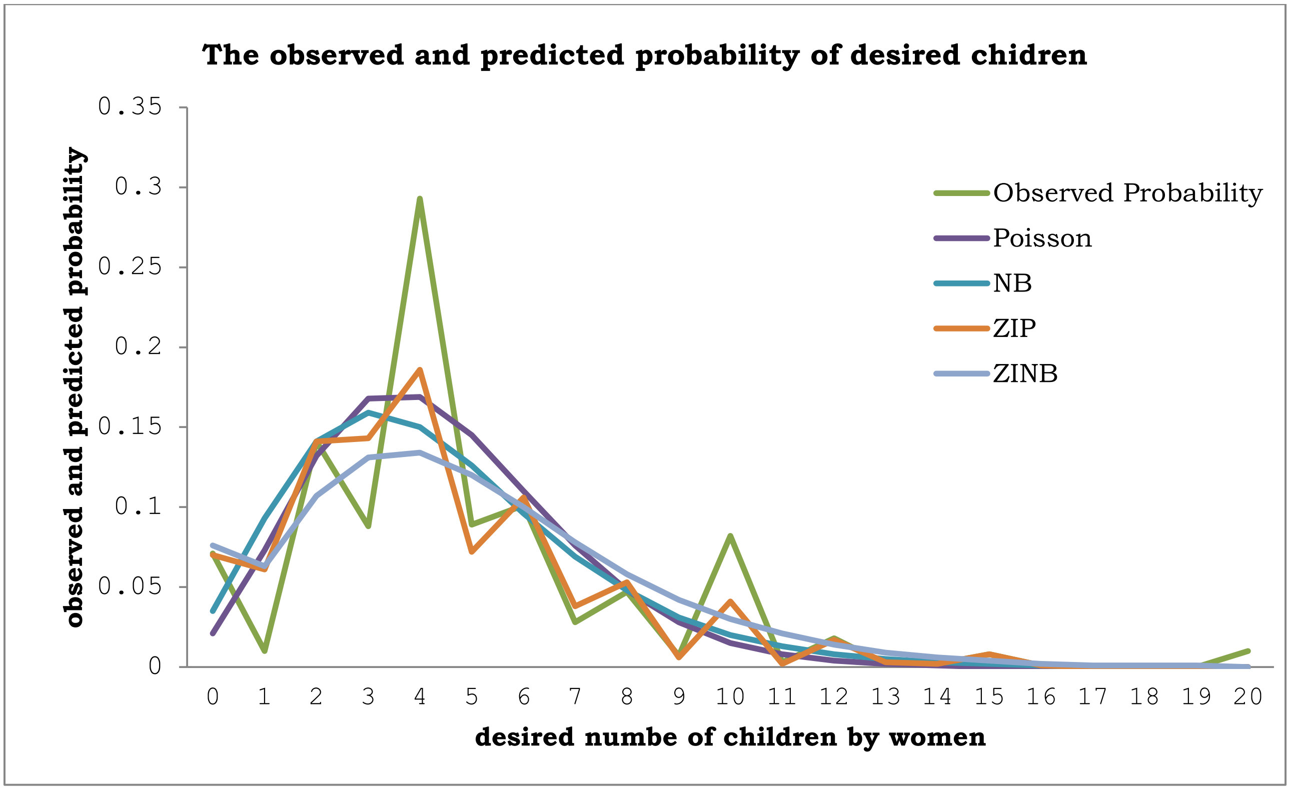
Appendix-2 Table 2.6. Shows the criteria to select the best model using without weighted data analysis. First, the calculated value of the Vuong test 20.73 was greater than the hypothetical value 1.96 for ZIP versus Poisson model. This value revealed that ZIP model was preferred to Poisson model in order to estimate the desired number of children among women. In the second case, comparison of ZINB versus NB models, the calculated value of the Vuong test is 20.280, revealed that the ZINB model was preferred to NB regression model. Finally, to compare the ZIP and ZINB models, using AIC and BIC and log likelihood the ZINB model is better fitted the desired number of children over lifetime of women data than did ZIP model. AIC and BIC values of ZINB was found to be small as compared to other count models in the case of the data set is without weighted. As a result of this the standard error of the selected count model ZINB is higher than that of the other models.

Table 4.8. Observed and predicted probabilities of desired number of children

		Probability				
Count	**Obs**	**Observed Prob.**	**Poisson**	**NB**	**ZIP**	**ZINB**
0	1,048	0.071	0.021	0.035	0.070	0.076
1	146	0.010	0.073	0.093	0.061	0.063
2	2,077	0.141	0.132	0.141	0.141	0.107
3	1,300	0.088	0.168	0.159	0.143	0.131
4	4,318	0.293	0.169	0.150	0.186	0.134
5	1,316	0.089	0.145	0.126	0.072	0.120
6	1,490	0.101	0.110	0.096	0.106	0.100
7	415	0.028	0.076	0.069	0.038	0.078
8	696	0.047	0.048	0.048	0.053	0.058
9	98	0.007	0.028	0.031	0.006	0.042
10	1,203	0.082	0.015	0.020	0.041	0.030
11	37	0.003	0.008	0.013	0.002	0.021
12	269	0.018	0.004	0.008	0.017	0.014
13	30	0.002	0.002	0.005	0.003	0.009
14	30	0.002	0.001	0.003	0.002	0.006
15	115	0.008	0.000	0.002	0.008	0.004
16	8	0.001	0.000	0.001	0.001	0.002
17	4	0.000	0.000	0.000	0.000	0.001
18	4	0.000	0.000	0.000	0.000	0.001
19	0	0.000	0.000	0.000	0.000	0.001
20	147	0.010	0.000	0.000	0.000	0.000

In this Table, weighted data set results indicates that predicted probabilities for ZIP model were the closest to the observed probabilities, the ZIP model is the most appropriate and preferred model than the other count models.

Figure4.2: The Observed and predicted probabilities of desired number of children



Since the predicted probabilities for ZIP model were the closest to the observed probabilities, the ZIP model is the most appropriate and preferred model than the other count models as shown in Table 4.8 and Figure 4.2.

From Figure 4.2 and the value of AIC, BIC and Vuong criteria in Table 4.7, it can be clearly observed that there were a difference among Poisson, NB, ZIP and ZINB models for the dataset. Therefore, it is possible to conclude that the ZIP model was more appropriate than the ZINB model to fit women desired number of children weighted dataset.

In the case of unweighted statistical data analysis the predicted probabilities for ZINB model were the closest to the observed probabilities, the ZINB model is the most appropriate and preferred model than the other count models as shown in Appendices-2 Table. A.2.5.

Table .4.9. The estimated Zero-inflated Poisson model for desired number of children of selected independent variables

parameters	Coef.	Robust Std. Err.	Z	P>|z|	IRR	[95% CI for IRR] Lower	Upper
women_age_5_year							
15-19	(Ref.)						
20-24	-0.163	0.019	-8.510	0.000	0.849	0.818	0.882
25-29	-0.298	0.021	-14.450	0.000	0.742	0.713	0.773
30-34	-0.341	0.024	-13.960	0.000	0.711	0.678	0.746
35-39	-0.403	0.024	-16.590	0.000	0.668	0.637	0.701
40-44	-0.495	0.028	-17.510	0.000	0.609	0.577	0.644
45-49	-0.553	0.030	-18.550	0.000	0.575	0.543	0.610
women_educ							
no education	(Ref.)						
Primary	-0.103	0.014	-7.320	0.000	0.902	0.877	0.927
Secondary	-0.188	0.024	-7.910	0.000	0.828	0.790	0.868
Higher	-0.229	0.028	-8.220	0.000	0.795	0.753	0.840
Place_Resid							
urban	(Ref.)						
Rural	0.055	0.026	2.120	0.034	1.057	1.004	1.112
Religion							
Orthodox	(Ref.)						
Catholic	0.066	0.048	1.380	0.168	1.069	0.972	1.174
Protestant	0.078	0.013	5.820	0.000	1.081	1.053	1.109
Muslim	0.250	0.014	17.490	0.000	1.284	1.248	1.320
Traditional	0.241	0.069	3.500	0.000	1.272	1.112	1.456
Other	0.195	0.055	3.560	0.000	1.215	1.092	1.353
wealth_index							
Poorest	(Ref.)						
Poorer	-0.060	0.018	-3.350	0.001	0.941	0.909	0.975
Middle	-0.044	0.019	-2.290	0.022	0.957	0.922	0.994
Richer	-0.063	0.019	-3.260	0.001	0.939	0.904	0.975
Richest	-0.086	0.029	-2.980	0.003	0.918	0.867	0.971
contra_use							
No	(Ref.)						
yes	-0.106	0.015	-7.260	0.000	0.899	0.874	0.925
women_occup							
no working	(Ref.)						
working	-0.056	0.012	-4.830	0.000	0.946	0.924	0.967
living_children							

has no living	(Ref.)						
has living children	0.135	0.017	7.760	0.000	1.145	1.106	1.184
_cons	-1.539	0.033	-46.400	0.000	0.215	0.201	0.229
ln(women_age)	1.000	(exposure)					
inflate							
women_age_5_year							
15-19	(Ref.)						
20-24	-0.541	0.265	-2.040	0.041	0.582	0.346	0.979
25-29	-0.053	0.225	-0.230	0.815	0.949	0.610	1.475
30-34	-0.181	0.238	-0.760	0.445	0.834	0.523	1.329
35-39	0.266	0.236	1.130	0.259	1.305	0.822	2.071
40-44	0.497	0.249	2.000	0.046	1.644	1.009	2.678
45-49	0.664	0.242	2.740	0.006	1.942	1.208	3.123
women_educ							
no education	(Ref.)						
Primary	-0.441	0.144	-3.050	0.002	0.644	0.485	0.854
Secondary	-2.354	1.647	-1.430	0.153	0.095	0.004	2.399
Higher	-15.976	0.827	-19.330	0.000	0.000	0.000	0.000
Place_Resid							
urban	(Ref.)						
Rural	0.679	0.274	2.480	0.013	1.972	1.153	3.374
Religion							
Orthodox	(Ref.)						
Catholic	0.296	0.433	0.680	0.494	1.345	0.575	3.143
Protestant	0.197	0.140	1.410	0.159	1.218	0.925	1.604
Muslim	0.310	0.130	2.400	0.017	1.364	1.058	1.758
Traditional	0.323	0.642	0.500	0.615	1.381	0.392	4.865
Other	0.756	0.374	2.020	0.043	2.130	1.024	4.429
wealth_index							
Poorest	(Ref.)						
Poorer	-0.250	0.166	-1.510	0.131	0.779	0.563	1.078
Middle	0.160	0.162	0.990	0.323	1.173	0.855	1.610
Richer	0.260	0.160	1.620	0.104	1.297	0.948	1.775
Richest	0.166	0.263	0.630	0.529	1.180	0.705	1.977
contra_use							
No	(Ref.)						
yes	-0.176	0.148	-1.190	0.235	0.839	0.627	1.121
women_occup							
no working	(Ref.)						
working	-0.082	0.113	-0.720	0.469	0.922	0.739	1.150
living_children							

has no living	(Ref.)						
has living children	0.059	0.190	0.310	0.757	1.060	0.731	1.539
_cons	-3.078	0.354	-8.700	0.000	0.046	0.023	0.092

$Statistically\ significant\ at\ \ 5\%, confidence\ level \quad p - value < 0.05$

4.2.3. Discussion and Interpretation of ZIP model fit results

In Table 4.9, estimated zero inflated Poisson regression model fit results of incident counts, the coefficients can be interpreted as follows: for a one unit change in the predictor variable, the log of the response variable is expected to change by the value of the regression coefficient (coef.). In ZIP model, for every one unit increase in a unit's of the significant predictors, the log number of desired children is expected to increase or decrease by approximately the corresponding coefficient in the column of coefficient (coef.). In this model the variables whose p-value<0.05, were considered statistically significant. To interpret the categorical weighted data made by the incidence rate ratios $\left(IRR = exp^{(coef.)}\right)$ which is important to explain the change in percentage $(IRR - 1)$ of significant predictors.

The top half of Table 4.9 contains coefficients for the factor change in the expected count for those in the Not Always-0 Group. The coefficients can be interpreted in the same way as coefficients from the PRM or the NBRM.

Women age: The estimated coefficients for all age groups of women are negative and statistically significant. The results in Table 4.9 show that age category of women has a significant impact on the desired number of children that women at reproductive age would like to have in the non-zero group. The expected desired number of children that women aged 20-24 would like to have in their lifetime was 0.849 times the expected desired number of children that women aged 15-19 would like to have in their lifetime holding all other variables in the model constant. In other words, the expected desired number of children that women aged 20-24 would like to have in their lifetime was 15.1% less than the expected desired number of children that women aged 15-19 would like to have.

That is, women aged 20-24 are expected to have 15.1% percent fewer desired number of children than women aged 15-19. Similarly, women aged *25-29, 30-34, 35-39, 40-44, 45-49* are expected to have 25.8%, 28.9%, 32.2%, 39.1%, and 42.5% respectively fewer desired number of children than women aged 15-19 holding all other variables in the model constant. Thus, the expected desired number of children decreases as the age group of women increases. The women age shows that decrease the proportion of desired number of children, when the age group of women is increase (See Table 4.1). Age differentials indicate highest proportion of desired number of children for the youngest age category (45-49 years) and has lowest proportion of desired children for the oldest age group (15-19 years), thus suggesting that the proportion of desired children fertility is negatively correlated with age.

That is, younger women have higher likelihood of having higher desired number of children. This result is consistent with the findings of Kabeer (2001) and Rahman (2012). A similar study also showed that the risk of exposure to wanted children and childbearing increases as age increases (Alemayehu et al 2010).

Living children

The positive and statistically significant value for *women having Living children* indicates that the desired number of children that women having living children would like to have in their lifetime is expected to be 14.5 percent more than the desired number of children that women with no children would like to have in their lifetime.

Women occupation: The estimated coefficient for women occupation is negative and statistically significant**.** The negative estimated coefficient, -0.057 for *working women* suggests that currently employed women are expected to have 5.4 percent fewer desired number of children than unemployed women(IRR=0.946). Working/employed women exhibited lower average desired number of children than unemployed/ not working women (see table 4.1).

Women education: In the fitted ZIP model, the coefficient for *primary education* is negative and statistically significant. Thus, primary level educated women are expected to have 9.8% percent fewer desired number of children than uneducated women (IRR=0.902). Similarly, the coefficients for both *secondary* and *higher* education are negative and statistically significant. The results reveal that secondary and higher educated women are expected to have 17.2 percent and 20.5 percent respectively fewer desired number of children than uneducated women controlling for other variables in the model. A similar study by Mokshed (2000) also showed that women having secondary or higher education were less likely to desire for additional children than those who had no education.

Place of residence: Desired number of children does differ by place of residence. The coefficient for *rural* is positive and statistically significant implying that the desired number of children that urban women would like to have is significantly different from the desired number of children that rural women would like to have.

This finding is contrary to the result of a study in Bangladesh by Rahman (2012) where it was shown that rural women desire to have more children than urban women (see Table 4.1). The difference in desire for additional children among rural and urban women could be due to the parents' perceived costs and benefits of children. Rural women perceive greater benefit from children and see lots of advantages in having large families.

Women Religion: The coefficients for the categories of religion: Protestant, Muslim, Traditional, Other are positive and statistically significant. The IRR imply that Protestant, Muslim, Traditional, and Other religion follower women are expected to have 8.1 percent, 28.4 percent, 27.2 percent, and 21.5 percent respectively higher desired number of children than Orthodox women.
There was no significant difference in the desired number of children that women would like to have between Catholic and orthodox women. These findings are similar to the findings of Alemayehu (2002).

Household Wealth index: The coefficients for Poorer, Middle, Richer, and Richest wealth index categories are negative and statistically significant. The IRRs imply that women in the poorer index are expected to have 5.9 percent less desired number of children than women in the poorest index. Likewise, women in the Middle, richer, and richest wealth index are expected to have 4.3 percent, 6.1 percent and 8.2 percent less desired number of children than women in the poorest wealth index. So, women in the highest wealth index are expected to have the lowest desired number of children among women in all wealth index categories, thus suggesting that desired children fertility is negatively correlated with wealth index (See table 4.2).

Contraceptive use: The negative and statistically significant coefficient, -0.106 suggest that the expected desired number of children that women currently using contraceptives would like to have in their lifetime is 10.1 percent less than the expected desired number of children that women not using contraceptives would like to have. Although it indicates that fertility of woman is negatively affected by severe economic crisis, the rate of responsiveness is not very high Women who ever used contraceptives are more likely to have five or more children than women who never used any contraceptives. The study by Ramesh (2010) showed that women who had ever used contraception had wanted more children than those who did not use contraception.

Interpretation of ZIP model for covariates of zero counts

The bottom half of Table 4.9, labeled “inflate”, contains coefficients for the factor change in the odds of being in the Always-0 Group compared to the Not Always-0 Group. As shown in Table 4.9, place of residence has a significant impact on the probability of being in the always zero group. The odds of being in the always zero group increased by 97.2% for rural women as compared to those women in urban centers controlling other variables in the model. Similarly, the odds of being in the always zero group increased by about 36% for Muslim women as compared to Orthodox women holding all other variables in the model constant. Conversely, the odds of being in the always zero group decreased by a factor of 0.643 for women with primary education as compared to women with no education holding all other variables in the model constant.

4.2.3. Comparison of precision of results: using weights vs. without sampling weights

The analyses based on the unweighted data are presented in Table 2.6 of appendix-2. The results showed that the ZINB is the most appropriate model. However, the precision of statistical results are not guaranteed because the data has not been adjusted by multiplier or sampling weights.

For instance, in the selected model standard error of the parameter estimation is higher than the other results of unweight data count models and the range of CI in the selected model is larger than the other count regression models.

Conversely, in the case of weighted data analysis, the standard error of the ML estimates of the selected model (ZIP) was lower than the standard error of the estimates for the other count models. Also the range of the CIs of the parameters of the selected model was smaller than range for the other count regression models.

Thus, the results of based on weighted data are more reliable than the results based on without weighted data. Since, a sampling error is usually measured in terms of the standard error and also CI for a particular statistic and also sampling errors are important data quality parameters which give a measure of the precision of the parameter estimates.

Generally, the results of weighted data analysis becomes too precise, reliable, and also guarantees unbiased estimates for population parameters and the standard error of the parameters must be minimum and similarly the range of CI must be smaller than the other count regression models.

5. Conclusions and Recommendations

5.1. Conclusions

This paper attempted to identify and analyze the determinants of the desired number of children that women at the reproductive age (15-49) would like to have in their lifetime using count regression model. Data from the 2011 Ethiopia Demographic and Health Survey (EDHS) were used for analysis. The use of the sampling weights when fitting models to such survey data has also been considered. The Zero-inflated Poisson regression model (ZIP) was found to be the most appropriate and preferred model among the four count models considered. The descriptive results suggested that there are high variability in the non-zero values. The variance of the desired number of 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 both NB and ZINB regression models.

Differences with respect to the best fitting count regression model were observed. While the ZIP model was found to be the best fitting model for the weighted sample data, ZINB was the best for the data without sample weights.
In order to make the results more precise, reliable, and guarantee unbiased estimates for population parameters, design weights should be taken into account.

This study captured predictor variables that had significant effects on the desired number of children.
The selected ZIP model fit results indicated that age of women, level of education of women, religion of women, household wealth index, contraceptive use, occupation of women, place of residence and whether a woman has living children were statistically significant factors influencing the desired number of children that Ethiopian women of reproductive age would like to have in their lifetime.

5.2. Recommendations

Based on the findings that we have obtained, we recommend the following issues:

1. *The government, concerned institutions and other involved stakeholders should consider the identified major factors while designing policy that will impact fertility decisions the most.*
2. *A further analysis taking in to account the points raised in the limitation and using post stratification weighting, as a strategy for correcting differences between sample and population parameter estimation are recommended.*

References

Ainsworth.M. Beegle.K,and Nyamete, A. (1996). *The impact of women's schooling on Fertility and contraceptive use*: a study of fourteen Sub-Saharan African countries. World Bank Economic Review, 10(1):85-122.

Abdul Hakim, (1994). *Factors affecting fertility in Pakistan*. The Pakistan Development Review33 (4): 685- 709.

Aklilu Z. W. (2011). *Use of Zero-Inflated Models for analysis of Immunological Data*. Inter university Institute for Biostatistics and Statistical Bioinformatics

Agresti, A. (2008). *An introduction to categorical data analysis second edition*. John Wiley and Son Inc., New York.

Agresti, A., (2002). *Categorical data analysis*. Second edition, John Wiley and Sons, Inc., Publication, New Jersey.

Ainsworth, M. (1989). *Socio-economic determinants of fertility in Coted'Ivoire*, LSMS Working paper No 53.World Bank, Washington DC.

Akaike, H. (1974). *A new look at the statistical model identification*. IEEE transactions on automatic control, 19(6): 716-723.

Akaike, H. (1973). "*Information Theory and an Extension of the Maximum Likelihood Principle*." Pp. 267–281 in Second International Symposium on Information Theory, edited by B. N. Petrov and F. Csaki. Budapest: Akademiai Kaido.

Akmam, W. (2002). Women's education and fertility rates in developing countries, with special reference to Bangladesh *Eubios Journal of Asian and International Bioethics* 12, 138-143.

Assefa H. (1992). *An Overview of the Determinants of High Fertility in Ethiopia*. Ethiopian Journal of Development Research, 14(2):1-30.

Atkins, D. C., & Gallop, R. J. (2007). *Rethinking how family researchers model infrequent outcomes:* A tutorial on count regression and zero-inflated models. Journal of Family Psychology, 21, 726–735.

Baldwin,C. S. (1977). *Catastrophe in Bangladesh*: an examination of alternative population growth possibilities 1975-2000.

Bankole A, Audam S. (2011). *Fertility preferences and contraceptive use among couples in sub-Saharan Africa*. African Popul Stud ;25(2):556–86.

Becker, G. S. (1960). *An economic analysis of fertility in demographic and economic change in developed countries.*

Becker, G. S. (1976). *The economic approach to human behavior*, the university of Chicago press.

Becker, G. S. and Lewis, H.G. (1973). *On the interaction between the quantity and quality of children*. Journal of political economy, 81(2 part 2): S279-S288.

Behrman, J.R. and Wolfe, B.L. (1984). *A more General approach to fertility determination in developing country.*

Casterline JB. *The pace of fertility transition:* national patterns in the second half of the twentieth century.

Cain M. (1982). *Perspectives on Family and Fertility in developing Countries*. Population Studies, 36(2):159-175.

Central Statistical Agency Ethiopia and ORC. Macro. (2011). *Ethiopia Demographic and Health Survey 2011/12.* Addis Ababa, and Calverton: Central Statistical Agency and ORC Macro.

Central Statistical Authority (2007). *The 2007 Population and Housing Census of Ethiopia*: Results at Country Level. Addis Ababa.

Central Statistical Authority (Ethiopia) and ORC Macro (2001). *Ethiopia Demographic and Health Survey (2000*). Addis Ababa and Calverton: Central Statistical Authority and ORC Macro.

Caldwell J.C., Caldwell P. (1987). The Cultural Context of High Fertility in sub-Saharan Africa. Population and Development Review, 16 (3), 409-437.

Cochrane S.H. (1979). *Fertility and Education.* Baltimore: The John Hopkins University Press.

Curling. T., Chen, L. C. and Hussein, S. B. (1976). *Demographic crisis*: the impact of the Bangladesh civil war (1971) on births and deaths in a rural area of Bangladesh.

Cox R. (1983). *Some Remarks on over dispersion.* Biometrical, 70: 269-274.

Easterlin, R. A. (1966). *On the relation of economic factors to recent and projected fertility changes.*

Dommaraju, P., and Agadjanian E.V. (2009) India's North-South divide and theories of fertility change. *Journal of Pop Research (2009)* 26: 249-72.

Easterlin, R. A. and Crimmins, E. M. (1985). *The fertility revolution, Supply-Demand Analysis.* University of Chicago Press, Chicago.

Easterlin, R. A., Pollak R. A. and Wachter, M.L. (1980). *Toward a more general model of fertility determination:* endogenous preferences and natural fertility. In: Easterlin RA (ed) population and economic change in developing countries.

Famoyes, F. (1993). *Restricted generalized Poisson Regression model.*

Communications in statistics-Theory and Methods.

Hilbe, J. M. (2011). *Negative binomial regression*, second edition, Cambridge University press.

Kabeer, N. (2001). *Ideas, economics and 'the sociology of supply':* explanations for fertility decline in b Bangladesh.

Kabir, A., Barbhuiya, M. A. K. and Islam, M. S. (2001). *Determinants of fertility in Bangladesh*: evidence from the three national surveys.

RanaEjaz Ali Khan, Khadija Malik Bari, (2014). *Revealing Ideal Number of Children by Women*: A Socioeconomic Analysis Using Pakistan Demographic and Health Survey.Pak J CommerSocSci Pakistan Journal of Commerce and Social Sciences 2014, Vol. 8 (3), 887-899

Kravdal, Ø. (2002). *Education and Fertility in sub-Saharan Africa:* individual and community effects.

Islam, S. (2009). *Differential determinants of birth spacing since marriage to first live birth in rural Bangladesh.* Pertanika Journal of social science and humanities 17(1):16

Leslie Kish, professor of sociology, program Director. (1965). *Survey Sampling*, Survey Research Center, Institute for social Research the University of Michigan.

Long J.S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. California: Sage Publications.

Long, J. S., &Freese, J. (2006). *Regression models for categorical dependent variables using stata (2nd, ed. ed.*). College Station, TX: Stata Press MacDevitt TM, Adlakha A, Fowler TB, Harris-Bourne V.

Diane Publication Company; (1996). *Trends in adolescent fertility and contraceptive use in the developing world.*

Melkersson, M and Rooth, D. (2000). *Modeling female fertility using inflated count data models.* Journal of population economics, 13(2): 189-203.

Miranda, A. (2010). *A double-hurdle count model for completed fertility data from the developing worl*. DoQSS working paper.

Mokshed, Ali. (2000). *The effect of selected socio-demographic characteristics on desire for additional children among couples in Bangladesh*. Mahidol University.

Mcquillan, K. (2004). *When does religion influence fertility?* Population and Development Review 30(1):25- 56.

Mason K.O. and Palan V.T. (1981). *Female employment and fertility in peninsular Malaysia:* The maternal role compatibility hypothesis reconsidered. Demography 18(4):549- 575.

Maudlin W.P.and Berelson B. (1978). *Conditions of fertility decline in developing countries, 1965- 1975*. Studies in Family Planning 9: 89- 145.

Mcquillan, K. (2004). *When does religion influence fertility?* Population and Development Review 30(1):25- 56.

Muhith, A. M. A. (2012). *Chronicle of Last Three Years*: Building the Future Budget Speech 2012-13, Minister, Ministry of Finance, Government of People's Republic of Bangladesh, Dhaka.

Mullahy, J. (1986). *Specification and some testing of modified count data models.* Journal of Econometrics, 33(3):341-365.

Nguyen-Dinh, H. (1997). *A socioeconomic analysis of the determinants of fertility:* the case of Vietnam. Journal of population economics, 10(3):251-271.

Nag, M. (1982). *Modernization and its impact on fertility:* the Indian Scene. The population council, New York.

Nagelkerke, N. J. D. (1991). *A note on a general definition of the coefficient of determination.* Biometrika, 78, 691- 692.

Osili, O. U. and Long. B. T. (2008). *Does female schooling reduce fertility?* Evidence from Nigeria. Journal of development economics, 87(1):57-75.

Pollak, R. A. and Watkins, S. C. (1993). *Cultural and economic approaches to fertility:* proper marriage or mésalliance? Population and development Review, 19(3):467-496.

Ramesh A. (2010). *Demographic, socio- economic and cultural factors affecting fertility differential in Nepal.* Geography and Population Department,Tribhuvan University, Kathmndul, Nepal.

Sano, Y., Jeong, Y., Acock, A. C., &Zvonkovic, A. M. (2005). *Working with count data: Practical demonstration of Poisson, negative binomial, and zero inflated regression models.* Proceedings of the 35th Annual Theory Construction and Research Methodology Workshop of the National Council of Family Relations, Phoenix, AZ.

Sennott, C., and Yeatman, S. (2012). Stability and change in fertility preferences among young women in Malawi.

Sidze et al, (ND). Changing fertility preferences in urban and rural Senegal: patterns and determinants.

Sharma S. (1998). *Relationship between gender of existing children and desire for additional children by Nepalese women.* Master's Thesis, Institute for Population and Social Research, Mahidol University.

Toefiqua Mahfuza. (2009). *Islam, influence of socio- demographic variables on fertility in Bangladesh.* Application of path model analysis.Pakistan journal of social science 6(5);313-320.Africa at the sub- national level.Institute of demography, university of Louvain Belgium.

Turkey PW. (1989). *Evolution and the Demand for Children.* Population and Development Review, 15(1), 61-90.

Vuong Q.H. (1989). *Likelihood ratio tests for model selection and non-nested hypotheses*. Econometrical, 57:2:307-333

Wang W, Felix F. (1997). *Modeling Household Fertility Decisions with Generalized Poisson Regression.* Journal of Population Economics, 10(3):273-283.

Winkelmann R, Zimmermann F. (1995). *Recent Developments in Count Data Modeling:* Theory and Application. Journal of Economic Surveys. 9(1):1-24.

Wang, W. and Famoyes, F. (1997). *Modeling household fertility decisions with generalized Poisson regression.* Journal of population economics, 10(3):273-283.

Winkelmann, R. (1995). *Duration dependence and dispersion in count-data models. Journal of Business and Economic Statistics,* 13(4):467-741.

Winkelmann, R. and Zimmermann, K. F. (1994). *Count data models for demographic data, Mathmatical Population Studies:* An International Journal of Mathematical Demography, 4(3):205-221.

Yaacob W. F. W., Lazim M. A. and Wah Y. B. (2010). *A Practical Approach in Modeling Count Data.* Proceedings of the Regional Conference on Statistical Sciences, 176-183.

Appendices-1. Analysis Using weighted data

Table .A.1.1. Estimates and standard errors for Poisson model

```
Poisson regression                              Number of obs    =      14751
                                                   Wald chi2(22)    =    1066.71
Log pseudolikelihood = -33768.068                  Prob > chi2      =     0.0000

--------------------------------------------------------------------------------------
                     |               Robust
                 INC |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+----------------------------------------------------------------
    women_age_5_year |
               20-24 |  -.1378347   .0211624    -6.51   0.000    -.1793123   -.0963571
               25-29 |  -.2988246   .0237331   -12.59   0.000    -.3453406   -.2523085
               30-34 |  -.3310309   .0276964   -11.95   0.000    -.3853148   -.2767469
               35-39 |  -.4305475   .0287423   -14.98   0.000    -.4868813   -.3742136
               40-44 |  -.5529294    .035117   -15.75   0.000    -.6217574   -.4841013
               45-49 |   -.632713   .0375691   -16.84   0.000     -.706347    -.559079
                     |
          women_educ |
             Primary |   -.066223   .0168288    -3.94   0.000    -.0992069   -.0332392
           Secondary |  -.1198176   .0253766    -4.72   0.000    -.1695547   -.0700804
              Higher |  -.1602456   .0297648    -5.38   0.000    -.2185835   -.1019076
                     |
         Place_Resid |
               Rural |   .0128864   .0311045     0.41   0.679    -.0480774    .0738501
                     |
            Religion |
            Catholic |   .0531896   .0569544     0.93   0.350    -.0584389    .1648181
          Protestant |   .0561555   .0163605     3.43   0.001     .0240896    .0882214
              Muslim |    .222602   .0173001    12.87   0.000     .1886945    .2565096
         Traditional |   .2113799   .0892376     2.37   0.018     .0364774    .3862825
               Other |   .0930178   .0820604     1.13   0.257    -.0678176    .2538533
                     |
        wealth_index |
              Poorer |  -.0443074   .0219898    -2.01   0.044    -.0874067   -.0012081
              Middle |  -.0617937   .0239482    -2.58   0.010    -.1087314   -.0148561
              Richer |  -.0923369   .0240295    -3.84   0.000    -.1394339     -.04524
             Richest |  -.1089868   .0354612    -3.07   0.002    -.1784895   -.0394841
                     |
          contra_use |
                 yes |   -.092742   .0176402    -5.26   0.000    -.1273161   -.0581678
                     |
         women_occup |
             working |  -.0519484   .0141368    -3.67   0.000    -.0796561   -.0242408
                     |
     living_children |
has living children  |   .1339239   .0191858     6.98   0.000     .0963205    .1715273
               _cons |  -1.578894   .0396067   -39.86   0.000    -1.656522   -1.501266
      ln(women_age)  |          1  (exposure)
-------------------------------------------------------------------------------------
```

```
. estat ic, n(14751)

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |  14751           .   -33768.07     23     67582.14    67756.92
-----------------------------------------------------------------------------
              Note:  N=14751 used in calculating BIC

. estat gof

         Deviance goodness-of-fit =  22996.29
         Prob > chi2(14728)       =    0.0000

         Pearson goodness-of-fit  =  19366.53
         Prob > chi2(14728)       =    0.0000


Poisson regression                                Number of obs   =      14751
                                                  Wald chi2(8)    =     738.54
Log pseudolikelihood = -34064.475                 Prob > chi2     =     0.0000

------------------------------------------------------------------------------
                 |               Robust
             INC |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
women_age_5_year |  -.1012883    .004954   -20.45   0.000    -.1109979   -.0915787
      women_educ |  -.0633668   .0098241    -6.45   0.000    -.0826218   -.0441119
     Place_Resid |   .0366508   .0217997     1.68   0.093    -.0060758    .0793774
        Religion |   .0015677   .0007905     1.98   0.047     .0000184    .0031169
    wealth_index |  -.0272949   .0068389    -3.99   0.000    -.0406989   -.0138908
      contra_use |  -.1147025   .0177626    -6.46   0.000    -.1495165   -.0798884
     women_occup |  -.0743946   .0141298    -5.27   0.000    -.1020884   -.0467008
 living_children |   .1208686   .0186169     6.49   0.000     .0843801    .1573572
           _cons |  -1.446328   .0584177   -24.76   0.000    -1.560824   -1.331831
  ln(women_age) |          1  (exposure)
------------------------------------------------------------------------------
```

Table.A.1.2. Estimates and standard errors for NB model

Negative binomial regression		Number of obs	= 14751
Dispersion	= mean	Wald chi2(22)	= 1079.57
Log pseudolikelihood	= -33274.915	Prob > chi2	= 0.0000

INC	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
women_age_5_year						
20-24	-.1347887	.0208724	-6.46	0.000	-.1756979	-.0938796
25-29	-.2983253	.0234708	-12.71	0.000	-.3443273	-.2523234
30-34	-.3325658	.0272963	-12.18	0.000	-.3860657	-.279066
35-39	-.4307927	.028468	-15.13	0.000	-.486589	-.3749964
40-44	-.5500921	.0348542	-15.78	0.000	-.6184051	-.4817791
45-49	-.6298215	.0373269	-16.87	0.000	-.7029809	-.5566622
women_educ						
Primary	-.0672024	.0165878	-4.05	0.000	-.0997138	-.034691
Secondary	-.1191246	.0252568	-4.72	0.000	-.168627	-.0696221
Higher	-.1613992	.0294706	-5.48	0.000	-.2191604	-.103638
Place_Resid						
Rural	.0174564	.0293272	0.60	0.552	-.0400239	.0749368
Religion						
Catholic	.0552505	.056261	0.98	0.326	-.055019	.16552
Protestant	.0624669	.0160055	3.90	0.000	.0310967	.093837
Muslim	.2250751	.017071	13.18	0.000	.1916167	.2585336
Traditional	.2176602	.0901148	2.42	0.016	.0410384	.394282
Other	.1034947	.0819241	1.26	0.206	-.0570736	.264063
wealth_index						
Poorer	-.0425774	.0215116	-1.98	0.048	-.0847393	-.0004155
Middle	-.0611419	.0234887	-2.60	0.009	-.1071789	-.0151049
Richer	-.0897104	.0235038	-3.82	0.000	-.135777	-.0436438
Richest	-.1043698	.0336099	-3.11	0.002	-.170244	-.0384956
contra_use						
yes	-.0915054	.0172186	-5.31	0.000	-.1252532	-.0577575
women_occup						
working	-.0473271	.0138031	-3.43	0.001	-.0743806	-.0202735
living_children						
has living children	.1317025	.0192036	6.86	0.000	.0940641	.1693409
_cons	-1.588116	.0380824	-41.70	0.000	-1.662757	-1.513476
ln(women_age)	1	(exposure)				
/lnalpha	-2.382753	.0861316			-2.551567	-2.213938
alpha	.0922962	.0079496			.0779594	.1092695

```
. estat ic, n(14751)

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |  14751   -34206.61   -33274.91     24     66597.83    66780.21
-----------------------------------------------------------------------------
               Note:  N=14751 used in calculating BIC


Negative binomial regression                    Number of obs    =      14751
Dispersion           = mean                     Wald chi2(8)     =     759.60
Log pseudolikelihood = -33481.401               Prob > chi2      =     0.0000

----------------------------------------------------------------------------------
                 |               Robust
             INC |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+----------------------------------------------------------------
women_age_5_year |  -.1015027   .0049126   -20.66   0.000    -.1111312   -.0918743
      women_educ |  -.0640011   .0096982    -6.60   0.000    -.0830091    -.044993
     Place_Resid |   .0388861   .0211583     1.84   0.066    -.0025835    .0803556
        Religion |   .0018094   .0008236     2.20   0.028     .0001953    .0034235
    wealth_index |  -.0263704   .0066659    -3.96   0.000    -.0394353   -.0133056
      contra_use |  -.1136803   .0174668    -6.51   0.000    -.1479147    -.079446
     women_occup |  -.0725364   .0138494    -5.24   0.000    -.0996807   -.0453921
 living_children |   .1180654   .0186383     6.33   0.000      .081535    .1545959
           _cons |   -1.45145   .0567088   -25.59   0.000    -1.562597   -1.340302
  ln(women_age)  |          1  (exposure)
-----------------+----------------------------------------------------------------
        /lnalpha |  -2.281225    .080669                      -2.439333   -2.123117
-----------------+----------------------------------------------------------------
           alpha |    .102159   .0082411                        .087219    .1196581
----------------------------------------------------------------------------------
```

Table. A.1.3. Estimates and standard errors for ZIP model

Zero-inflated Poisson regression		Number of obs =	14751
		Nonzero obs =	13703
		Zero obs =	1048
Inflation model = logit		Wald chi2(22) =	1512.26
Log pseudolikelihood = -31507.04		Prob > chi2 =	0.0000

INC	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
INC						
women_age_5_year						
20-24	-.1632313	.0191734	-8.51	0.000	-.2008104	-.1256522
25-29	-.2978453	.0206067	-14.45	0.000	-.3382338	-.2574569
30-34	-.340623	.0243961	-13.96	0.000	-.3884385	-.2928076
35-39	-.4029171	.0242847	-16.59	0.000	-.4505142	-.3553201
40-44	-.4952223	.0282821	-17.51	0.000	-.5506543	-.4397903
45-49	-.5525567	.0297901	-18.55	0.000	-.6109442	-.4941693
women_educ						
Primary	-.1031017	.0140907	-7.32	0.000	-.1307191	-.0754844
Secondary	-.1884114	.0238248	-7.91	0.000	-.235107	-.1417157
Higher	-.2293219	.0278874	-8.22	0.000	-.2839802	-.1746635
Place_Resid						
Rural	.0551215	.0259588	2.12	0.034	.0042431	.1059999
Religion						
Catholic	.0664404	.0481439	1.38	0.168	-.0279199	.1608008
Protestant	.0775661	.0133214	5.82	0.000	.0514567	.1036755
Muslim	.2497077	.0142803	17.49	0.000	.2217189	.2776966
Traditional	.2407281	.0688108	3.50	0.000	.1058614	.3755949
Other	.194864	.0547123	3.56	0.000	.0876298	.3020981
wealth_index						
Poorer	-.0602822	.0180141	-3.35	0.001	-.0955893	-.0249752
Middle	-.0440522	.0192031	-2.29	0.022	-.0816897	-.0064147
Richer	-.0630603	.0193276	-3.26	0.001	-.1009417	-.0251789
Richest	-.0858692	.028777	-2.98	0.003	-.142271	-.0294674
contra_use						
yes	-.1064661	.0146558	-7.26	0.000	-.1351911	-.0777412
women_occup						
working	-.055875	.0115666	-4.83	0.000	-.078545	-.0332049
living_children						
has living children	.1350561	.0174119	7.76	0.000	.1009294	.1691828
_cons	-1.53903	.0331671	-46.40	0.000	-1.604036	-1.474024
ln(women_age)	1	(exposure)				
inflate						
women_age_5_year						
20-24	-.5414695	.2652374	-2.04	0.041	-1.061325	-.0216138
25-29	-.0526657	.2251539	-0.23	0.815	-.4939593	.3886278
30-34	-.1814707	.2377149	-0.76	0.445	-.6473835	.284442
35-39	.2662733	.2356758	1.13	0.259	-.1956428	.7281895

```
             40-44  |   .4970091   .2489442     2.00   0.046     .0090874    .9849307
             45-49  |   .6637849   .2423833     2.74   0.006     .1887224    1.138847
                    |
        women_educ  |
           Primary  |  -.4406957   .1444915    -3.05   0.002    -.7238939   -.1574975
         Secondary  |  -2.353994   1.647449    -1.43   0.153    -5.582936     .874947
            Higher  |   -15.9762    .826563   -19.33   0.000    -17.59623   -14.35616
                    |
       Place_Resid  |
             Rural  |   .6791549   .2739196     2.48   0.013     .1422824    1.216027
                    |
          Religion  |
          Catholic  |   .2961972   .4331273     0.68   0.494    -.5527168    1.145111
        Protestant  |   .1974744   .1403609     1.41   0.159    -.0776278    .4725767
            Muslim  |   .3104862   .1295405     2.40   0.017     .0565914    .5643809
       Traditional  |   .3229012   .6424567     0.50   0.615    -.9362908    1.582093
             Other  |    .755903   .3736017     2.02   0.043      .023657    1.488149
                    |
      wealth_index  |
            Poorer  |  -.2497719   .1655299    -1.51   0.131    -.5742047    .0746608
            Middle  |   .1596213   .1615304     0.99   0.323    -.1569724     .476215
            Richer  |   .2600876    .160202     1.62   0.104    -.0539026    .5740778
           Richest  |   .1657901   .2632655     0.63   0.529    -.3502008    .6817809
                    |
        contra_use  |
               yes  |  -.1760358   .1481361    -1.19   0.235    -.4663773    .1143056
                    |
       women_occup  |
           working  |  -.0815486   .1127453    -0.72   0.469    -.3025252    .1394281
                    |
   living_children  |
has living children |   .0587122   .1901371     0.31   0.757    -.3139498    .4313741
             _cons  |  -3.078102    .353874    -8.70   0.000    -3.771682   -2.384522
-------------------------------------------------------------------------------------

. estat ic, n(14751)

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |  14751   -32897.87   -31507.04     46     63106.08    63455.64
-----------------------------------------------------------------------------
               Note:  N=14751 used in calculating BIC
```

Table. A.1.4. Estimates and standard errors for ZINB model

```
Zero-inflated negative binomial regression        Number of obs    =      14751
                                                  Nonzero obs      =      13703
                                                  Zero obs         =       1048

Inflation model      = logit                      Wald chi2(22)    =    1484.00
Log pseudolikelihood = - 31507.04                  Prob > chi2      =     0.0000

------------------------------------------------------------------------------------
                     |               Robust
                 INC |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+--------------------------------------------------------------
INC                  |
    women_age_5_year |
              20-24  |  -.1566213   .0188155    -8.32   0.000     -.193499   -.1197436
              25-29  |  -.2937722   .0204642   -14.36   0.000    -.3338813   -.2536631
              30-34  |  -.3349499   .0242905   -13.79   0.000    -.3825585   -.2873414
              35-39  |  -.3983331   .0241782   -16.47   0.000    -.4457215   -.3509446
              40-44  |  -.4903586   .0281438   -17.42   0.000    -.5455193   -.4351978
              45-49  |  -.5482589   .0296682   -18.48   0.000    -.6064075   -.4901103
                     |
          women_educ |
            Primary  |  -.1057232   .0141325    -7.48   0.000    -.1334224   -.0780241
          Secondary  |  -.1831031   .0228186    -8.02   0.000    -.2278266   -.1383795
             Higher  |  -.2261529   .0276661    -8.17   0.000    -.2803775   -.1719284
                     |
         Place_Resid |
              Rural  |   .0573345    .026024     2.20   0.028     .0063284    .1083406
                     |
            Religion |
           Catholic  |   .0626826   .0486195     1.29   0.197    -.0326099    .1579752
         Protestant  |   .0765516   .0132881     5.76   0.000     .0505074    .1025958
             Muslim  |   .2477254   .0143417    17.27   0.000     .2196162    .2758346
        Traditional  |   .2381596   .0692495     3.44   0.001      .102433    .3738862
              Other  |   .1939146   .0547679     3.54   0.000     .0865714    .3012578
                     |
         wealth_index |
             Poorer  |   -.057243   .0179648    -3.19   0.001    -.0924533   -.0220326
             Middle  |   -.043266   .0191846    -2.26   0.024    -.0808671   -.0056648
             Richer  |  -.0625306   .0194031    -3.22   0.001    -.1005599   -.0245012
            Richest  |  -.0828038   .0288555    -2.87   0.004    -.1393595   -.0262481
                     |
          contra_use |
                yes  |  -.1058648   .0146441    -7.23   0.000    -.1345666   -.0771629
                     |
         women_occup |
            working  |  -.0567889   .0115726    -4.91   0.000    -.0794709   -.0341069
                     |
     living_children |
has living children  |   .1345881   .0173361     7.76   0.000       .10061    .1685661
               _cons |  -1.544497   .0333898   -46.26   0.000     -1.60994   -1.479054
      ln(women_age)  |          1  (exposure)
---------------------+--------------------------------------------------------------
inflate              |
    women_age_5_year |   .1677677   .0337465     4.97   0.000     .1016258    .2339097
          women_educ |  -.6108871   .1143516    -5.34   0.000    -.8350122    -.386762
         Place_Resid |   .9353233   .2260919     4.14   0.000     .4921912    1.378455
            Religion |   .0074773    .003653     2.05   0.041     .0003176     .014637
        wealth_index |   .1089541   .0486802     2.24   0.025     .0135426    .2043657
          contra_use |  -.2192852   .1458158    -1.50   0.133    -.5050789    .0665085
```

women_occup	-.1377813	.1082966	-1.27	0.203	-.3500388	.0744762
living_children	-.1722644	.1642245	-1.05	0.294	-.4941386	.1496098
_cons	-4.662396	.531111	-8.78	0.000	-5.703355	-3.621438
/lnalpha	-9.988062	77.81599	-0.13	0.898	-162.5046	142.5285
alpha	.0000459	.0035753			2.66e-71	7.93e+61

```
. estat ic, n(14751)
```

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	14751	-32805.07	-31546.24	33	63108.08	63465.25

Note: N=14751 used in calculating BIC

```
Zero-inflated negative binomial regression        Number of obs   =      14751
                                                  Nonzero obs     =      13703
                                                  Zero obs        =       1048

Inflation model      = logit                      Wald chi2(8)    =    1021.24
Log pseudolikelihood = -31891.25                  Prob > chi2     =     0.0000
```

INC	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
INC						
women_age_5_year	-.0837949	.00396	-21.16	0.000	-.0915563	-.0760334
women_educ	-.0930392	.0088024	-10.57	0.000	-.1102915	-.0757868
Place_Resid	.0882341	.0185298	4.76	0.000	.0519163	.1245519
Religion	.0026263	.0005653	4.65	0.000	.0015183	.0037342
wealth_index	-.0181059	.0056366	-3.21	0.001	-.0291534	-.0070584
contra_use	-.1310769	.0149066	-8.79	0.000	-.1602933	-.1018604
women_occup	-.0851367	.0117527	-7.24	0.000	-.1081717	-.0621018
living_children	.1052495	.0166636	6.32	0.000	.0725895	.1379096
_cons	-1.508867	.0493874	-30.55	0.000	-1.605665	-1.41207
ln(women_age)	1	(exposure)				
inflate						
women_age_5_year	.1714272	.0338923	5.06	0.000	.1049996	.2378548
women_educ	-.6265926	.1154669	-5.43	0.000	-.8529036	-.4002815
Place_Resid	.9110297	.2278279	4.00	0.000	.4644952	1.357564
Religion	.0071262	.0037564	1.90	0.058	-.0002362	.0144887
wealth_index	.1065188	.0489169	2.18	0.029	.0106434	.2023943
contra_use	-.2120366	.1463025	-1.45	0.147	-.4987844	.0747111
women_occup	-.1499793	.1086632	-1.38	0.168	-.3629553	.0629967
living_children	-.1476095	.1676458	-0.88	0.379	-.4761893	.1809703
_cons	-4.640746	.5362696	-8.65	0.000	-5.691815	-3.589677
/lnalpha	-4.803798	.5578695	-8.61	0.000	-5.897202	-3.710393
alpha	.0081986	.0045737			.0027471	.0244679

Appendices-2. Analysis without Using weighted data

Table .A.2.1. Estimates and standard errors for Poisson model

Poisson regression

Number of obs	=	14751
LR chi2(22)	=	6099.03
Prob > chi2	=	0.0000
Pseudo R2	=	0.0783

Log likelihood = -35902.54

INC	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
women_age_5_year						
20-24	-.1462409	.0142646	-10.25	0.000	-.174199	-.1182829
25-29	-.3129791	.0155168	-20.17	0.000	-.3433915	-.2825667
30-34	-.3712372	.0169462	-21.91	0.000	-.4044512	-.3380232
35-39	-.4411195	.0171201	-25.77	0.000	-.4746743	-.4075647
40-44	-.5258894	.0184336	-28.53	0.000	-.5620185	-.4897603
45-49	-.6449162	.0193892	-33.26	0.000	-.6829184	-.606914
women_educ						
Primary	-.1313043	.0099143	-13.24	0.000	-.1507359	-.1118727
Secondary	-.1990079	.0177737	-11.20	0.000	-.2338436	-.1641721
Higher	-.2304903	.0212289	-10.86	0.000	-.2720982	-.1888825
Place_Resid						
Rural	-.0155385	.0158404	-0.98	0.327	-.0465852	.0155081
Religion						
Catholic	-.0162561	.0382474	-0.43	0.671	-.0912196	.0587073
Protestant	.1023525	.0112059	9.13	0.000	.0803893	.1243157
Muslim	.3167765	.0091125	34.76	0.000	.2989163	.3346367
Traditional	.1396394	.0501558	2.78	0.005	.0413358	.2379429
Other	.2360343	.0377375	6.25	0.000	.1620702	.3099984
wealth_index						
Poorer	-.1676171	.0121932	-13.75	0.000	-.1915153	-.143719
Middle	-.1862294	.012519	-14.88	0.000	-.2107661	-.1616927
Richer	-.1834339	.0122355	-14.99	0.000	-.2074151	-.1594527
Richest	-.2526828	.0171038	-14.77	0.000	-.2862057	-.21916
contra_use						
yes	-.1138038	.0108972	-10.44	0.000	-.135162	-.0924456
women_occup						
working	-.0782608	.0078645	-9.95	0.000	-.093675	-.0628465
living_children						
has living children	.1162645	.0127313	9.13	0.000	.0913117	.1412173
_cons	-1.345657	.0209379	-64.27	0.000	-1.386694	-1.30462
ln(women_age)	1	(exposure)				

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	14751	-38952.06	-35902.54	23	71851.08	72025.86

Note: N=14751 used in calculating BIC

```
. estat gof

         Deviance goodness-of-fit =  25442.69
         Prob > chi2(14728)       =    0.0000

         Pearson goodness-of-fit  =  22698.78
         Prob > chi2(14728)       =    0.0000
```

Table.A.2.2. Estimates and standard errors for NB model

```
Negative binomial regression                      Number of obs   =      14751
                                                  LR chi2(22)     =    3283.62
Dispersion     = mean                             Prob > chi2     =     0.0000
Log likelihood = -34812.923                       Pseudo R2       =     0.0450
```

INC	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
women_age_5_year						
20-24	-.1440454	.0176107	-8.18	0.000	-.1785618	-.109529
25-29	-.3140748	.0194681	-16.13	0.000	-.3522316	-.2759181
30-34	-.3718023	.02155	-17.25	0.000	-.4140396	-.329565
35-39	-.4397135	.0219769	-20.01	0.000	-.4827874	-.3966396
40-44	-.5233265	.0239316	-21.87	0.000	-.5702316	-.4764214
45-49	-.640234	.0252326	-25.37	0.000	-.689689	-.5907789
women_educ						
Primary	-.1284282	.0125884	-10.20	0.000	-.153101	-.1037554
Secondary	-.1924476	.0218225	-8.82	0.000	-.2352189	-.1496762
Higher	-.223506	.0258254	-8.65	0.000	-.2741229	-.1728892
Place_Resid						
Rural	-.006008	.0197729	-0.30	0.761	-.0447622	.0327461
Religion						
Catholic	-.0090175	.0476968	-0.19	0.850	-.1025015	.0844665
Protestant	.106438	.014116	7.54	0.000	.0787712	.1341048
Muslim	.3200774	.0116925	27.37	0.000	.2971606	.3429942
Traditional	.1444527	.0656204	2.20	0.028	.015839	.2730664
Other	.2286121	.0510213	4.48	0.000	.1286122	.328612
wealth_index						
Poorer	-.1633513	.0160135	-10.20	0.000	-.1947372	-.1319654
Middle	-.1845498	.0163685	-11.27	0.000	-.2166314	-.1524682
Richer	-.1812759	.015993	-11.33	0.000	-.2126215	-.1499302
Richest	-.2502161	.0216822	-11.54	0.000	-.2927124	-.2077197
contra_use						
yes	-.1101323	.0137174	-8.03	0.000	-.1370179	-.0832467

```
      women_occup |
          working  |  -.0700132   .0101416    -6.90   0.000    -.0898904    -.050136
                   |
  living_children |
has living children  |   .1161765   .0160358     7.24   0.000      .084747     .147606
            _cons |  -1.362887   .0263205   -51.78   0.000    -1.414474     -1.3113
   ln(women_age) |          1  (exposure)
-------------------+----------------------------------------------------------------
         /lnalpha |  -2.073651   .0312871                      -2.134973   -2.012329
-------------------+----------------------------------------------------------------
            alpha |   .1257259   .0039336                       .1182478     .133677
------------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0:  chibar2 (01) = 2179.23 Prob>=chibar2 = 0.000
```

```
Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |  14751   -36454.73   -34812.92     24     69673.85    69856.22
-----------------------------------------------------------------------------
               Note:  N=14751 used in calculating BIC
```

Table. A.2.3. Estimates and standard errors for ZIP model

```
Zero-inflated Poisson regression                  Number of obs    =      14751
                                                  Nonzero obs      =      13703
                                                  Zero obs         =       1048

Inflation model = logit                           LR chi2(22)      =    6344.34
Log likelihood  = -33683.98                       Prob > chi2      =     0.0000

------------------------------------------------------------------------------------
                INC |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------------+----------------------------------------------------------------
INC                |
    women_age_5_year |
             20-24  |  -.1560609   .0144487   -10.80   0.000    -.1843798    -.127742
             25-29  |  -.3096172   .0157177   -19.70   0.000    -.3404233    -.278811
             30-34  |  -.3632776   .0171351   -21.20   0.000    -.3968617   -.3296934
             35-39  |  -.4189373   .0173167   -24.19   0.000    -.4528775   -.3849971
             40-44  |  -.4784814   .0186398   -25.67   0.000    -.5150148    -.441948
             45-49  |  -.6020563   .0195329   -30.82   0.000    -.6403402   -.5637724
                   |
       women_educ |
          Primary  |  -.1553871   .0101151   -15.36   0.000    -.1752123   -.1355618
        Secondary  |  -.2503048   .0180894   -13.84   0.000    -.2857593   -.2148503
           Higher  |  -.2877705    .021444   -13.42   0.000       -.3298    -.245741
                   |
      Place_Resid |
            Rural  |  -.0099482   .0159175    -0.62   0.532     -.041146    .0212496
                   |
         Religion |
         Catholic  |   .0270068   .0392363     0.69   0.491    -.0498949    .1039086
       Protestant  |    .110627   .0113734     9.73   0.000     .0883354    .1329185
           Muslim  |   .3319484    .009263    35.84   0.000     .3137932    .3501036
```

```
        Traditional  |   .1589763   .0508359     3.13   0.002     .0593397    .2586129
              Other  |   .2714784   .0378526     7.17   0.000     .1972887    .3456682
                     |
        wealth_index |
             Poorer  |  -.1579903   .0123238   -12.82   0.000    -.1821446   -.1338361
             Middle  |  -.1623288   .0127045   -12.78   0.000    -.1872292   -.1374284
             Richer  |  -.1707803   .0124016   -13.77   0.000    -.1950871   -.1464736
            Richest  |  -.2518281   .0172318   -14.61   0.000    -.2856017   -.2180544
                     |
          contra_use |
                yes  |  -.1239562   .0110443   -11.22   0.000    -.1456027   -.1023098
                     |
         women_occup |
            working  |  -.0840092   .0080009   -10.50   0.000    -.0996908   -.0683276
                     |
     living_children |
has living children  |   .1165277   .0128991     9.03   0.000     .0912458    .1418095
               _cons |  -1.288642   .0212142   -60.74   0.000    -1.330221   -1.247063
      ln(women_age) |          1  (exposure)
---------------------+----------------------------------------------------------------
inflate              |
    women_age_5_year |   .1210267   .0239051     5.06   0.000     .0741736    .1678798
          women_educ |  -.6689773    .077337    -8.65   0.000    -.8205549   -.5173996
            Religion |   .0030606   .0033587     0.91   0.362    -.0035223    .0096434
        wealth_index |  -.0278817   .0271053    -1.03   0.304    -.0810072    .0252438
          contra_use |   -.193646   .1140544    -1.70   0.090    -.4171885    .0298965
         women_occup |  -.1791646   .0764219    -2.34   0.019    -.3289488   -.0293805
     living_children |  -.1158881   .1195569    -0.97   0.332    -.3502154    .1184391
               _cons |  -2.544543   .1360059   -18.71   0.000     -2.81111   -2.277976
--------------------------------------------------------------------------------------
Vuong test of zip vs. standard Poisson:            z =    20.73  Pr>z = 0.0000

. estat ic, n(14751)

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |  14751   -36856.15   -33683.98     31     67429.96    67665.53
-----------------------------------------------------------------------------
               Note:  N=14751 used in calculating BIC
```

Table. A.2.4. Estimates and standard errors for ZINB model

```
Zero-inflated negative binomial regression       Number of obs   =      14751
                                                 Nonzero obs     =      13703
                                                 Zero obs        =       1048

Inflation model = logit                          LR chi2(22)     =    4425.77
Log likelihood  = -33505.76                      Prob > chi2     =     0.0000

------------------------------------------------------------------------------------
                 INC |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+--------------------------------------------------------------
INC                  |
    women_age_5_year |
              20-24  |  -.1520338    .015689    -9.69   0.000    -.1827838   -.1212839
              25-29  |  -.3078843    .017207   -17.89   0.000    -.3416094   -.2741592
              30-34  |  -.3608094    .018881   -19.11   0.000    -.3978156   -.3238033
              35-39  |  -.4147821   .0191836   -21.62   0.000    -.4523812    -.377183
              40-44  |  -.4737408   .0207939   -22.78   0.000     -.514496   -.4329856
              45-49  |  -.5951168    .021823   -27.27   0.000     -.637889   -.5523445
                     |
          women_educ |
            Primary  |  -.1569487   .0111617   -14.06   0.000    -.1788251   -.1350723
          Secondary  |  -.2473379    .019532   -12.66   0.000    -.2856198   -.2090559
             Higher  |  -.2830674   .0230564   -12.28   0.000     -.328257   -.2378777
                     |
         Place_Resid |
              Rural  |  -.0055775   .0174144    -0.32   0.749    -.0397091    .0285541
                     |
            Religion |
           Catholic  |   .0272857   .0430156     0.63   0.526    -.0570234    .1115948
         Protestant  |   .1127816   .0124746     9.04   0.000      .088332    .1372313
             Muslim  |   .3343877   .0102487    32.63   0.000     .3143007    .3544747
        Traditional  |   .1603504   .0569673     2.81   0.005     .0486967    .2720042
              Other  |   .2682275   .0431185     6.22   0.000     .1837167    .3527383
                     |
        wealth_index |
             Poorer  |  -.1571747   .0138161   -11.38   0.000    -.1842537   -.1300957
             Middle  |   -.164019   .0142272   -11.53   0.000    -.1919038   -.1361341
             Richer  |  -.1713578   .0138782   -12.35   0.000    -.1985586   -.1441569
            Richest  |  -.2512839       .019   -13.23   0.000    -.2885231   -.2140446
                     |
          contra_use |
                yes  |  -.1229091   .0121241   -10.14   0.000    -.1466719   -.0991463
                     |
         women_occup |
            working  |  -.0809801   .0088903    -9.11   0.000    -.0984047   -.0635556
                     |
     living_children |
has living children  |   .1167722   .0141591     8.25   0.000     .0890208    .1445236
               _cons |  -1.300526   .0232864   -55.85   0.000    -1.346167   -1.254886
      ln(women_age)  |          1  (exposure)
---------------------+--------------------------------------------------------------
inflate              |
    women_age_5_year |    .133142   .0248445     5.36   0.000     .0844477    .1818363
          women_educ |  -.7320231    .085392    -8.57   0.000    -.8993882   -.5646579
            Religion |   .0034929   .0033827     1.03   0.302    -.0031371    .0101228
        wealth_index |  -.0257299   .0285026    -0.90   0.367    -.0815941    .0301342
          contra_use |  -.2254291   .1215364    -1.85   0.064     -.463636    .0127778
```

women_occup	-.1901225	.0801045	-2.37	0.018	-.3471243	-.0331206
living_children	-.0764678	.1294863	-0.59	0.555	-.3302563	.1773208
_cons	-2.659665	.1482007	-17.95	0.000	-2.950133	-2.369197
/lnalpha	-3.229269	.063628	-50.75	0.000	-3.353977	-3.10456
alpha	.0395864	.0025188			.0349451	.0448442

Vuong test of zinb vs. standard negative binomial: z = 20.28 Pr>z = 0.0000

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	14751	-35718.64	-33505.76	32	67075.52	67318.69

Note: N=14751 used in calculating BIC

Table. A.2.5. Observed and predicted probability count models

		Probability				
Count	**Frequency**	**Observed Probability**	**Poisson**	**NB**	**ZIP**	**ZINB**
0	1,048	0.071	0.019	0.035	0.075	0.073
1	146	0.010	0.064	0.088	0.055	0.063
2	2,077	0.141	0.116	0.128	0.100	0.127
3	1,300	0.088	0.149	0.143	0.131	0.131
4	4,318	0.293	0.155	0.136	0.138	0.184
5	1,316	0.089	0.139	0.117	0.127	0.100
6	1,490	0.101	0.112	0.094	0.106	0.100
7	415	0.028	0.084	0.073	0.083	0.028
8	696	0.047	0.060	0.054	0.061	0.058
9	98	0.007	0.040	0.039	0.044	0.042
10	1,203	0.082	0.026	0.028	0.030	0.030
11	37	0.003	0.016	0.020	0.020	0.021
12	269	0.018	0.009	0.014	0.013	0.014
13	30	0.002	0.005	0.009	0.008	0.009
14	30	0.002	0.003	0.006	0.005	0.006
15	115	0.008	0.002	0.004	0.003	0.004
16	8	0.001	0.001	0.003	0.002	0.002
17	4	0.000	0.000	0.002	0.001	0.001
18	4	0.000	0.000	0.001	0.000	0.001
20	147	0.010	0.000	0.001	0.000	0.000

Table. A.2.6. comparison of model used in without weighted data

Selection Criteria	Models			
	Poisson	**NB**	**ZIP**	**ZINB**
Log liklihood	-35902.54	-34812.92	-33683.98	-33505.76
AIC	71851.08	69673.85	67429.96	67075.52
BIC	72025.86	69856.22	67665.53	67318.69
Vuong			20.73	20.28
(p-value)			0.00	0.00