College of Natural and Computational Sciences
Department of Statistics

Trends and Determinants of Teenage Fertility in Ethiopia

By: Sosna Bayu

A Thesis Submitted to the Department of Statistics in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics (Biostatistcs)

Advisor: Mekonnen Tadesse(Assoc. Prof)

June, 2020

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 material used for the thesis have been duly acknowledged.

Sosna Bayu	. . . .	. . . .
Name	Signature	Date

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

Mekonnen Tadesse (Assoc. Prof)	. . . .	. . . .
Name	Signature	Date

Addis Ababa University
College of Natural and Computational Sciences
Department of Statistics

Trends and Determinants of Teenage of Fertility in Ethiopia

By
Sosna Bayu

As members of the Board of Examiners of M.Sc. thesis open defense examination of the above title, we have evaluated the thesis and examined the candidate.

Eshetu Wencheko (Prof.)	....	....
Examiner	Signature	Date
Dejen Tesfaw (PhD)	....	....
Examiner	Signature	Date
Merga Belina (PhD)	....	....
Chairman, Department graduate committee	Signature	Date

Acknowledgment

First, and foremost, I would like to extend my unshared thanks to almighty God and his mother saint Mary.

I would like to express my special thanks to my advisor Mekonnen Tadesse (Assoc. Prof.), for his commitment, guidance and constructive suggestions. My special thanks also go to Mr.Zelalem Tazu and all staff members of the Department of Statistics for the encouragements and support they provided for me.

I would also thank all members of my family for all the encouragement and support throughout my study. My appreciation also goes to my friends, classmates and for all who gave me any support to complete this thesis.

Finally, I would like to thank the Central Statistical Agency for providing the EDHS data.

Abstract

Trends and Determinants of Teenage of Fertility in Ethiopia
Sosna Bayu
Addis Ababa University, 2020

Fertility during adolescence is associated with increased risk of maternal mortality, morbidity and premature termination of education. Once teenagers become pregnant or have children, they face serious problems which have impact on their family and the society at large. Data from four Ethiopia demographic and health surveys (EDHSs) conducted in 2000, 2005, 2011 and 2016 were used to describe the trend of teenage fertility in Ethiopia. Moreover, the 2016 EDHS data on women aged 15-19 years have been used to identify factors associated with teenage fertility. Multilevel (three-level) logistic regression models were used not only to identify the major risk factors of teenage fertility but also to determine variations of teenage fertility among various EAs and HH using the 2016 EDHS data. The findings from the multilevel logistic regression showed that religion, wealth status and ever-been married are significantly associated with teenage fertility in Ethiopia.

Contents

List of Tables

List of Figures

Acronyms

WHO World Health Organization

UN United Nations

PBR Population Reference Bureau

STDs Sexually Transmitted Diseases

HIV Human Immune-deficiency Virus

AIDS Acquired immunodeficiency syndrome

ICPD International Conference on Population and Development

SRH Sexual Reproductive Health

EDHS Ethiopia Demographic and Health Survey

CSA Central Statistical Agency

PHC Population and Housing Census

EA's Enumeration Areas

ICC Intra-Class Correlation Coefficient

HH Household

PSU Primary sampling unit

1 Introduction

1.1 Background

Adolescence is a period of transition from childhood to adulthood characterized by significant physiological, psychological and social changes. World health organization has classified the age groups 10-19 and 15-24 years as adolescents and youth, respectively. Those in the age group 10-24 years are called young people. Adolescence is a period of key changes in life within a short period of time. Chief among these transitions for many girls is that they are exposed to sex, marriage and pressured to conceive after marriage (UN, 2019).

In demography, fertility indicates the product or output of reproduction instead of ability to have children. The physiological ability to have children that is manifest roughly in the period between menarche and menopause in women is termed fecundity. Demographers have outlined a third additional side of reproduction called fecundability which is the probability of becoming pregnant, or the chance of exposure to the possibility that depends on the pattern of sexual and pregnancy preventive behaviors (Frank, 2017).

Adolescent fertility refers to fertility among 15 to 19 year olds. Early child bearing expose adolescent women to unnecessary risks. Their chance of dying is twice as high as that of a woman who waited until her 20s to start childbearing. In addition, early childbearing greatly reduces the chance of a girl advancing her education and limits her opportunities for training and employment. Globally, the trend has decreased, however there are very certain variations in levels and trends across regions (PBR, 2013).

Adolescent mothers are more likely to suffer severe complications throughout delivery and have higher levels of morbidity and mortality. Compared to older mothers, adolescent mothers are less likely to utilize health care services and they are more likely to initiate antenatal care late, deliver outside health facilities and rely on an unskilled birth attendants (WHO, 2012). Considerations regarding adolescent childbearing include its association with heightened health risks for mothers and their infants in addition to lower educational attainment and persistent poverty among women who become mothers through out early adolescence (UN, 2019). Giving birth as a teenager leads to having more children than women who start childbearing after their teenage years. The outcome of such phenomenon contributes to high population growth. Teenage pregnancy is also associated with social stigma, stillbirth, low birth weight and maternal death.

Society shapes key values that dictate when young women begin bearing children and how their biological process is regarded by their families in addition by health and welfare services. For very young married women living in rural areas, wherever society might believe early childbearing as normal and even fascinating, the social and economic risks of not bearing children probably outweigh the physical risks of bearing children. In contrast to remote rural areas, urban areas have higher health care access, not to mention long-term educational and training opportunities factors that ought to, in theory, improve health outcomes for young women and their children (National academic press, 1993).

Adolescent fertility worldwide over the past twenty years, since the landmark Programme of Action of the 1994 International Conference on Population and Development, is turning into less common in almost every country. However, several countries are still characterized by slow declines or high levels of early childbearing. More investments for keeping girls in school, delaying the age at wedding and increasing access to information and services to prevent pregnancy will help ensure that all adolescents will fulfill their right to work out if, when and how many children they want to have (UN, 2013).

The level of fertility in the world varies broadly by country and culture, social and economic conditions, as well as by individual characteristics such as age (Frank, 2017). The incidence of adolescent maternity and adolescent birth is widely divergent. The highest adolescent birth rates occur in Sub-Saharan Africa (143 per 1000) and in some countries in South Asia and Latin America (156 per 1000) (WHO, 2004). In developing countries, complications throughout pregnancy and childbirth are main causes of death among adolescent girls. The risk is highest for girls below 15 years of age (UN, 2019).

Adolescent childbearing encompasses a negative impact on the health of the adolescents and their infants; individual social and economic effects; and social level impacts. Studies have shown rates of newborn death to average about 50% higher to adolescent mothers as compared with mothers in their 20s (Macro International 2008).

Even though maternal and newborn health programs have a clear role in better serving the wants of the teenage mothers, additional and thorough analysis is necessary to address the evidence base on effective interventions for pregnant adolescents and to translate knowledge into action at country level.
This study will therefore, attempt to identify and explore factors associated with teenage fertility and study trends of teenage fertility in Ethiopia.

1.2 Statement of the problem

Fertility is the main element to have an effect on the health of the mother. Adolescent fertility has increasingly been viewed as a source of social and policy concern. Fertility before a girl is physically, developmentally and socially ready threatens her right to a successful transition into adulthood. Teenagers are not prepared to give birth and to support their babies since they are young intellectually and economically. This results in adolescent mothers as well as their children suffer from adverse health consequences. Teenagers will face different health problems and children of teens also get insufficient health care and receive inadequate parenting that will affect late the future of nations. In most studies, despite the fact that, teenage fertility varies among the national population, girls living in developing countries are the most in danger of adolescent fertility.

Many adolescents lack basic information regarding the symptoms, transmission, and treatment of STDs exposing adolescents to a larger risk of getting disease. Very young mothers are most at risk of fistula. They additionally suffer from greater risks of sepsis as a result of sophisticated deliveries. Most reproductive health problems experienced by adolescents are also experienced by older women. But they are exacerbated among the young, whether or not by physical immatureness, primiparity, or social condemnation. (National academic press, (1993))

Teenage fertility is associated with, high risk of maternal and child morbidity and mortality, and other social consequences. Also, it is a leading cause of death for adolescent girls between the ages of 15 and 19 in poor countries. Ethiopia located in Sub-Saharan Africa is one country where teenage fertility is a problem facing adolescent girls. Therefore, reducing teenage fertility is an obvious place to anchor serious efforts to reduce poverty and to create in future generations. This study is intended to determine the determinant of teenage fertility in Ethiopia and its association with different factors. It also studies the trends in teenage fertility over four different periods of demographic and health national surveys.

1.3 Objective of the study

The aim of this study is to assess the trend and determinants of teen-age fertility in Ethiopia. An attempt will be made to assess the risk factors that can lead to teenage pregnancy and birth.

1.4 Significance of the study

Reproductive behavior has changed substantially and certain original simplifying assumptions have become less accurate over time. Concerns associated with teenage fertility varies from time to time and, therefore an updated study is necessary. Even though adolescent fertility is a very sensitive issue, studies on it are very limited in our country. This in turn leads to a lack of awareness and gives rise to misunderstandings about the consequences. Therefore, this study was conducted with the aim of assessing the determinant factors associated with adolescent fertility so as to contribute to the prevention programs by providing up-to-date information for decision-making and program implementation. The trends of teenage fertility needs due attention; as well.

Addressing the multiple factors and reducing adolescent fertility are essential for improving sexual and reproductive health and, ultimately, the social and economic wellbeing of adolescent girls. A good understanding of teenage fertility can result in improved services to this very vulnerable special group. There are various critical reasons why it is necessary to understand and learn a lot of concerning teenage fertility. The context of fertility for adolescents becomes crucial because of their susceptibility to early sexual intercourse and childbearing, that subsequently accelerates the potential risk of maternal mortality and morbidity.
In summary, the results of our study are expected

- to expand the evidence base on teenage fertility by providing the findings of this study.
- to help not only in laying out an action plan or policy that can address adolescent pregnancy but also in improving Maternal and Newborn Survival.
- to increase the understanding of teenage pregnancy in Ethiopia

1.5 Limitations of the study

- There are various factors that are not included in the survey like rape, family experience of adolescent birth and peer pressure. If this factors is considered in this study, it may affect teenage fertility in Ethiopia.
- The data don't show the current situation.

2 Review of related Literature

Several studies concerning teenage fertility and pregnancy are conducted across the world. Teenage fertility can occur in all societies, with considerable variation in magnitude and consequences among different countries and regions. Most of the studies and reports indicate that teenage fertility is higher in Africa and particularly in Sub-Saharan Africa. WHO (2019) reported that in 2016, maternal mortality was the second main cause of death for women of reproductive age, after HIV/AIDS, and was the main cause among women aged 15-29 years and additionally the majority maternal deaths (95%) occurred in low-income and lower-middle-income countries, and almost two thirds (65%) occurred within the World Health Organization (WHO) African Region.

Worldwide, an estimated 16 million girls between 15 and 19 years old give birth each year and two million girls under age fifteen, become pregnant every year. Cross-national comparative statistics cover up significant differences and disparities within countries. In all countries, the poorest, least educated girls are far more likely to become adolescent mothers as compared to advantaged peers. This difference take poor girls of their rights (Blum et al., 2015). According to the report by WHO (2008), Ethiopia is one of the seven countries where half of all adolescent births occur. In the poorest regions of the world, birth rates for 15-19 year olds are still four times higher than in the high-income regions. About 90% of adolescent births among 15-19 year olds occur within marriage.

According to the report by WHO (2019), in 2018, there were an estimated 12.8 million births among adolescent girls aged 15-19 years, representing 44 births per 1000 adolescent girls. The lowest adolescent birth rates occur in high income countries and highest in low income countries. Regionally, adolescent birth rates are lowest and highest in the WHO Western Pacific Region (14 births per 1000) and within the African Region (99 births per 1000) respectively. This report additionally show that there is an inverse association between adolescent birth rates and the proportion of women aged 15-49, years with their contraceptive desires met, that is additionally associated with the wealth of countries. The adolescent birth rates and the proportion of women receiving quality care from a skilled health professional are also inversely associated. Thus, not only do adolescents giving birth have less access to ways of preventing risky births, but also risky births are less likely to be delivered by professional birth attendant.

According to the UN (2019) report, globally the adolescent fertility rate declined from 56 births per 1,000 adolescent girls in 2000 to 45 in 2015 and 44 in 2018. However, the adolescent fertility rate remained high in sub-Saharan Africa at 101 births per 1,000 adolescent girls in 2018.

According to a discussion paper on adolescence by WHO (2019), even though sexual activity starts at an early age in several developed countries in Europe, adequate contraceptive method is usually available for adolescents, as are abortion facilities in several countries. Despite variation between such countries, the number of pregnancies and births among adolescents is relatively low compared to other regions of the world. The woman's worry of losing her partner or incurring his anger seem to be an important factor that discourage young females from exercising choice in the timing of sexual activity or negotiating for the utilization of condoms or alternative contraceptives. In many parts of the world, particularly in sub-Saharan Africa, South Asia and Latin America, unsafe abortion is a commonly occurred and a reason of illness and even death, particularly among young girls.

According to a report by UN on Adolescent Fertility at the 2013 International Conference on Population and Development (ICPD) in Cairo, the cumulative global population of female adolescents living in countries with a low or medium adolescent fertility level where the average birth rate is less than 80 births per 1,000 for women aged 15-19, increased from 62 percent in 1990-1995 to 83 percent in 2005-2010. The evidence presented also shows strong associations between all measures of different aspects of education in a country and adolescent fertility.

A study on early fertility and parenthood in the transition to adulthood by Swartz et al. (2018) indicates that having a child at an early age puts huge pressure on young people, their relationships with their life partners and on their relations to help to take care the needs of the child. Even though there is an access to the alternative opportunities during child bearing, it is harder to achieve. Pregnancy and fertility for teenagers are more about developing a sense of adult self-worth as they are about having a child.

The finding of the study from Nepal on trends and factors associated with pregnancies among adolescent women by Poudel et al. (2018) suggested that adolescent pregnancy was significantly higher among woman with middle household wealth index or poor household wealth index and for unemployed women. In contrast, adolescent pregnancy was significantly lower among women with access to media and exposure to public health issues.

A study conducted by Kassa et al. (2018) on prevalence and determinants of adolescent pregnancy in Africa, using a systematic review and meta-analysis of published and unpublished studies in Africa, revealed that the overall pooled prevalence of adolescent pregnancy in Africa was 18.8% while it was 19.3% in the Sub-Saharan African region. The prevalence was highest (21.5%) in Eastern Africa and lowest (9.2%) in Northern Africa . According to this study, factors associated with adolescent pregnancy include place of residence, ever-married, school attendance, maternal education, fathers education and lack of parent to adolescent communication on sexual and reproductive health (SRH) issues.

The findings of a study from Nigeria on utilization of maternal health care services among married adolescent by Raiesh et al. (2012) suggested that preceding and current maternal health protection programmatic efforts have not accomplish in reaching adolescent women, serving their specific needs or addressing the socioeconomic barrier that lie in the way of their utilizing existing maternity services. Over two-thirds of adolescent women were younger than 18 years of age at the time of childbearing. The majority of adolescent women were residing in rural areas and every three in five had no formal education. More than half of the adolescents were not working and about two-fifths had no mass media exposure.

A study conducted in Northern Ghana by Yussif et al. (2017) on long-term effects of adolescent pregnancies on subsequent pregnancies and births of the young mothers, revealed that as compared to older women, adolescent women have 80% higher risk for a cesarean section for the first two successive births. The findings of this study have shown that pregnant teenagers experience more abortions, cesarean sections and stillbirths in their lifetime which expose them to death and other diseases. Adolescent and older mothers included in the study had a very low socioeconomic status.

Alemayehu et al. (2010) studied adolescent fertility in Ethiopia using data from EDHS conducted in 2005. They applied multivariate logistic regression model to identify sociodemographic and economic determinants of adolescent fertility. Bongaarts model was also used to find out proximate determinants of fertility. Their findings showed that women living in rural areas were more likely to be fertile than their counterparts living in the urban areas. Adolescents living in rural and other urban areas outside Addis Ababa were four and two times more likely to be fertile than adolescents living in Addis Ababa, respectively. The overall teenage fertility shows that one out of six female adolescents aged between 15 and 19 years were either pregnant or had already given birth. Teenage fertility rate in Ethiopia is relatively lower than the rates documented for Kenya, Uganda, Tanzania, and Malawi which were but higher than some African countries like Eritrea and Rwanda. The differences noted may be ascribed to the variations in demographic and socio-cultural conditions of the countries such as marital age, age at first sexual intercourse and use of family planning. In this study woman education and place of residence were found to be strong determinants of adolescent fertility.

A multilevel analysis of the predictors of teenage pregnancy in Ethiopia by Birhanu et al. (2019) identified many factors at both the individual and community level using data from the 2016 Ethiopian Demographic and Health Survey. Their study revealed that about 79.6% of women aged 20-24 years experienced pregnancy throughout their adolescent stage. Being sexually active before age 15; being married before age 15(individual level) and being a rural dweller (community level) show positive association with teenage pregnancy. Communities with a lower proportion of contraceptive users had also a statistically significant association with teenage pregnancy.

Cross-sectional study on teenage pregnancy and its associated factors by Mathewos et al. (2018) used multivariate logistic regression to identify predictors of teenage pregnancy of high school adolescents in Arba-minch town and marital status, religious involvement, awareness of fertile period of menstrual cycle and educational status of the respondents mothers did not show significant associations with teenage pregnancy. In this study, the risk of teenage pregnancy was higher to those who were living with both biological parents.

3 Data and Methodology

3.1 Data Source

The data for this study were extracted from four Ethiopia Demographic and Health Surveys (EDHS) which have been conducted by the Central Statistical Agency (CSA). We used all four EDHS 2000, 2005, 2011 data for trend analysis and only the 2016 EDHS data to assess the determinants of teenage fertility. The 2016 EDHS is the fourth Demographic and Health Survey conducted in Ethiopia. The survey was carried out from January 18 to June 27, 2016, based on a nationally representative sample that provides estimates at the national and regional levels and for urban and rural areas. The main objective of the 2016 EDHS was to provide up-to-date estimates of key demographic and health measures. The EDHS provides an entire summary of population, maternal, and child health issues in Ethiopia.

The sampling frame used for the 2016 EDHS is the Ethiopia Population and Housing Census (PHC) that was conducted in 2007 by the Central Statistical Agency of Ethiopia. The census frame is a complete list of 84,915 enumeration areas (EAs) created for the 2007 PHC. The administrative regions in Ethiopia are divided into 11 geographical regions. Each region is sub-divided into zones, each zone into woredas, each woreda into towns, and every town into kebeles. The sample for the 2016 EDHS was designed to provide estimates of populations and health indicators for the country as a whole, for urban and rural areas separately, and for each of the nine regions and the two administrative cities. The 2016 EDHS sample was stratified and selected in two stages. Each region was stratified into urban and rural areas, giving rise to 21 sampling strata. An enumeration area is a geographic area consisting of a convenient number of dwelling units that served as a counting unit for the census. Samples of EAs were selected independently in each stratum in two stages.

In the initial stage 645 enumeration areas were selected. Some of the chosen EAs were large, consisting over 300 households. To minimize the task of household listing, each large EA selected for the 2016 EDHS was segmented. Just one segment was selected for the survey with a probability proportional to segment size. Household listing was conducted only within the selected segment; that is, a 2016 EDHS cluster is either an EA or a segment of an EA.

In the second stage of selection, a fixed number of 28 households per cluster (enumeration area) were selected with an equal probability systematic selection from the newly created household listing. The survey interviewer interviewed only the pre-selected households. No replacements or changes of the pre-selected households was allowed in the implementing stages to prevent bias. All females aged 15-49 and all males aged 15-59 who were either permanent residents of the selected households or visitors who stayed in the household the night before the survey were eligible to be interviewed.

3.2 Variables of the Study

3.2.1 The Response Variable

The response variable in this study is teenage fertility excluding women aged 20 and above. So if a teenager is either currently pregnant or has already given birth, the outcome is coded as 1 and if not it is coded as 0. The response variable for the ith subject in jth2nd level and in the kth 3rd cluster is represented by y$_{ijk}$ with two possible values 0 and 1. In EDHS data, enumeration areas (PSUs) representing blocks of HHs are at level 3, households at level 2 and individuals at level 1.

3.2.2 Explanatory Variables

In this study, the socio economic and demographic characteristics included as predictor variables are: wealth index (HH level), level of education, religion, region, media exposure, place of residence (EA level), currently working status, ever-been married, and current contraceptive use. The descriptions of these variables are presented below (Table 1).

Table 1: Descriptions of the predictor variables considered in the study

No	Variable	Categories
1	Ever-been married	(0)No (1)Yes
2	Wealth index	(0)Poor (1)Middle (2)Rich
3	Level of education	(0)No education (1)Primary (2)Secondary or Above
4	Religion	(0) Orthodox (1)Catholic (2)Protestant (3)Muslim (4)Others
5	Currently working	(0)No (1)Yes
6	Media Exposure	(0)Have no exposure (1)have exposure
7	Region	(1)Tigray (2)Afar (3)Amhara (4)Oromia (5)Somilia (6)Benishangul-Gumuz (7)SNNP (8)Gambela (9)Harari (10)Addis Ababa (11)DireDewa
8	Place of residence	(0)Urban (1)Rural
9	Current Contraceptive use	(0)No (1)Yes

3.3 Methodology

3.3.1 Multilevel models

Multilevel analysis is a methodology for the analysis of data with complex structure of variability associated with nested sources of variability. Multilevel analysis is used when the data have clustered structures, with elementary units at level 1 nested within clusters at level 2, which successively nested within clusters at level 3, and so on. The multilevel logistic regression analysis looks at variations due to hierarchy structure in the data. It allows the simultaneous examination of the effects of group level (cluster and division) and individual level variables on individual level outcomes while accounting for the non-independence of observations within-groups. The analysis also allows the examination of both between-group and within-group variability as well as how group level and individual level variables are related to variability at both levels.

Test of heterogeneity Proportions

For the proper application of multilevel analysis, the first logical step is to test heterogeneity of proportions between groups. The most commonly used test statistic to check for heterogeneity of proportions between groups is the Chi-square test. The Chi-square test is given as follows:

$$X^2 = \sum_{j=1}^{N} n_j \frac{(\hat{p}_j - \hat{p}_.)}{\hat{p}_.(1-\hat{p}_.)} \tag{1}$$

where, $\hat{p}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} y_{ij}$ is the proportion of successes in group j which is an estimate for the group-dependent probability P_j. Similarly, $\hat{p}_. = \frac{1}{M}\sum_{j=1}^{N}\sum_{i=1}^{n_i} Y_{ij}$ the overall proportion of successes. N is the number of groups, n_j is the number of samples in the j^{th}group. The decision were based on approximately Chi-square distribution with N-1 degrees of freedom.

In EDHS data, enumeration areas (PSUs) representing blocks of HHs are at level 3, households at level 2 and individuals at level 1. Consider a dichotomous outcome variable $y_{ijk} \sim$ Bernoulli (π_{ijk}), then the logit link function is:

$$\eta_{ijk} = \beta_0 + \beta_1 x_{ijk} + \beta_2 x_{HH(jk)} + \beta_3 x_{PSU(k)} + \varepsilon_{j(k)}(2) + \varepsilon_k^{(3)} \tag{2}$$

where $\eta_{ijk} = ln(\frac{\pi_{ijk}}{1-\pi_{ijk}})$, π_{ijk} denotes the probability that the i^{th} teenager in the j-th 2nd level cluster and the k-th 3rd level cluster gives birth or is pregnant; i.e. $\pi_{ijk} = P(Y_{ijk} = 1)$, x_{ijk} denotes the vector of the teenager-level variables, $x_{HH(jk)}$ denotes the vector of the second level variables, and $x_{PSU(k)}$ denotes the vector of 3rd level predictor variables. In addition, β_1 denotes the vector of regression parameters for the teenager-level variables, β_2 denotes the vector of regression parameters for second level variables, and β_3 denotes the vector of regression parameters for the 3rd level variables. $\varepsilon_{j(k)}^{(2)}$ denotes the random effect for the j-th level cluster in the k-th level cluster and $\varepsilon_k^{(3)}$ denotes the random effect for k-th 3rd level cluster. We assume that $\varepsilon_{j(k)}^{(2)} \sim N(0, \sigma_{HH}^2)$ and $\varepsilon_k^{(3)} \sim N(0, \sigma_{PSU}^2)$ are independent.

Intercept-only model

This is the simplest case of a hierarchical model for a dichotomous outcome variable in which there are no explanatory variables at all. This model only contains random groups and random variation within groups.

$$\eta_{ijk} = \beta_0 + \varepsilon_{j(k)}^{(2)} + \varepsilon_k^{(3)} \tag{3}$$

where $\varepsilon_{j(k)}^{(2)}$ denotes the random effect for the j-th level cluster in the k-th level cluster and $\varepsilon_k^{(3)}$ denotes the random effect for k-th 3rd level cluster.

Random intercept model

In this model the covariates are included and the intercept is the only random effect meaning that the groups differ with respect to the average value of the response variable. This model assume the slopes are fixed.

$$\eta_{ijk} = \beta_0 + \beta_1 x_{ijk} + \beta_2 x_{HH(jk)} + \beta_3 x_{PSU(k)} + \varepsilon_{j(k)}(2) + \varepsilon_k^{(3)} \tag{4}$$

where x_{ijk} denotes the vector of the teenager-level variables, $x_{HH(jk)}$ denotes the vector of the second level variables, and $x_{PSU(k)}$ denotes the vector of 3rd level predictor variables. In addition, β_1 denotes the vector of regression parameters for the teenager-level variables, β_2 denotes the vector of regression parameters for second level variables, and β_3 denotes the vector of regression parameters for the 3rd level variables. $\varepsilon_{j(k)}^{(2)}$ denotes the random effect for the j-th level cluster in the k-th level cluster and $\varepsilon_k^{(3)}$ denotes the random effect for k-th 3rd level cluster.

Random Coefficient Model

In this model the coefficients of the explanatory variables are considered as random. Random coefficients models treat covariate as well as the intercept as random variables that can explain unobserved heterogeneity in the effects of explanatory variables on the response variable.

$$\eta_{ijk} = \beta_0 + \beta_1 x_{ijk} + \beta_2 x_{HH(jk)} + \beta_3 x_{PSU(k)} + \varepsilon_{j(k)}^{(2)} + \varepsilon_{oj(k)}^{(2)} x_{ijk} + \varepsilon_k^{(3)} + \varepsilon_{ok}^{(3)} x_{ijk} \tag{5}$$

where x_{ijk} denotes the vector of the teenager-level variables, $x_{HH(jk)}$ denotes the vector of the second level variables, and $x_{PSU(k)}$ denotes the vector of 3rd level predictor variables. In addition, β_1 denotes the vector of regression parameters for the teenager-level variables, β_2 denotes the vector of regression parameters for second level variables, and β_3 denotes the vector of regression parameters for the 3rd level variables. $\varepsilon_{j(k)}^{(2)}$ denotes the random effect for the j-th level cluster in the k-th level cluster and $\varepsilon_k^{(3)}$ denotes the random effect for k-th 3rd level cluster. $\varepsilon_{oj(k)}^{(2)}$ and $\varepsilon_{ok}^{(3)}$ denote random slope. The part $\beta_0 + \beta_1 x_{ijk} + \beta_2 x_{HH(jk)} + \beta_3 x_{PSU(k)}$ of equation (5) is called fixed part of the model and $\varepsilon_{j(k)}^{(2)} + \varepsilon_{oj(k)}^{(2)} x_{ijk} + \varepsilon_k^{(3)} + \varepsilon_{ok}^{(3)} x_{ijk}$ is random part of the model.

3.3.2 Parameter Estimation

The standard multilevel modeling approach can properly estimate parameters and standard errors for clustered data that have resulted from equal probability sampling. However, with unequal probabilities of selection the standard Maximum likelihood method does not estimate the model parameters and their standard errors appropriately. So we consider sampling weight in estimating the parameter to take in to account the unequal probability selection. Pseudo maximum likelihood estimation is used for estimating the parameter. Survey designs often involve unequal selection probabilities of clusters and/or people within clusters. This survey includes design (sampling) weights to account for unequal selection probabilities. When estimating multilevel models that are based on survey data, sampling weights are incorporated into the likelihood, producing a pseudo likelihood.

Weighting scheme

Due to the non-proportional allocation of the sample to different regions and their urban and rural areas and, possible differences in response rates, a sampling weight must be used in the analyses of the 2016 EDHS data to ensure the actual representativeness of the survey results. The sampling weights are based on sampling probabilities separately for each sampling stage and each cluster. We use the following notations:
$P1h_i$: initial-stage sampling probability of the ith cluster in stratum h
$P2h_i$: 2^{nd} stage sampling probability within the ith cluster (households).
Let a_h be the number of EAs selected in stratum h, M_{hi} the number of households according to the sampling frame in the i^{th} EA, and $\sum Mhi$ the total number of households in the stratum. The probability of selecting the i^{th} EA in the 2016 EDHS sample is calculated as:

$$\frac{a_h M_{hi}}{\sum Mhi} \ .$$

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

$$P_{1h_i} = \frac{a_h M_{hi}}{\sum Mhi} \times z_{hi} \ .$$

Let L_{hi} be the number of households 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 selection probability for a household in the cluster is calculated as:
$P_{2hi} = \frac{g_{hi}}{L_{hi}}$
The overall selection probability of a household in cluster i of stratum h is the product of the selection probabilities of the two stages:

$$P_{hi} = P_{1hi} \times P_{2hi} \ .$$

The sampling weight for each household in cluster i of stratum h is the inverse of its overall selection probability:

$$w_{hi} = \frac{1}{p_{hi}} \ .$$

The sampling weights are incorporated into the likelihood for estimation of parameters and their standard errors (CSA, 2016). We maximize the weighted likelihood with respect to the regression parameters.
Let y_{ijk} be the response vector then the full log-pseudo likelihood is:

$$L_w(y) = \sum w_{hi} L_{whi}(y_{ijk}). \tag{6}$$

Scaling the weights

After creating the dataset, one needs to scale the weights. In any analysis, here or otherwise, one should scale the weights before doing any other data management, as the scaled weight should be based on the entire sample of individuals with weights.

$$w_{ij}^* = w_{ij} \left[\frac{\sum_i w_{ij}}{\sum_i w_{ij}^2} \right]$$

where
w_{ij}^* represents the scaled weight for individual i in cluster j and w_{ij} unscaled weight for individual i in cluster j (Carle, 2009).

3.3.3 Intraclass Correlation

Individual units in a cluster often behave more alike than units from different clusters. There exists a correlation between observations when they belong to the same cluster. The amount of cluster and household variation is expressed as Intra-class Correlation Coefficient(ICC). For three level binary data the ICC is often defined for each level separately:

$$ICC_{PSU} = \frac{\sigma_{PSU}^2}{\sigma_{PSU}^2 + \sigma_{HH}^2 + \frac{\pi^2}{3}} ----------- ICCattributabletolevel3$$
$$ICC_{HH} = \frac{\sigma_{HH}^2}{\sigma_{PSU}^2 + \sigma_{HH}^2 + \frac{\pi^2}{3}} -------------- ICCattributabletolevel2$$

where $\frac{\pi^2}{3}$ denotes the variation of lower (individual) level unit, σ_{PSU}^2 and σ_{HH}^2 denote the variation between EAs and households, respectively.

3.3.4 Assessing Goodness of Fit

It is necessary to check how well the model fits the data after a model has already been fitted. Therefore, goodness of fit of the model can be assessed in two ways: first by testing the overall goodness of the model and second by testing the significance of each predictor variable in the model.

Likelihood Ratio Test

Overall fit of a model can be tested by comparing the fit of the models with and without the independent variables. A model with k independent variables (the given model) is claimed to provide a better fit to the data, if it demonstrates an improvement over the model with no independent variables (the null model). The overall fit of the model with all coefficients in the model can be examined through a likelihood ratio test which tests the null hypothesis $H_0 : \beta_1 = \beta_2 = ... = \beta_k = 0$. To do this, the deviance with just the intercept (-2 log likelihood of the null model) is compared to the deviance when the k independent variables have been added (-2 log likelihood of the given model). The likelihood of the null model is the likelihood of obtaining the observations if the independent variables have no effect on the outcome where the likelihood of the given model is the likelihood of obtaining the observations with all independent variables incorporated in the model. The difference of these two yields a goodness of fit index statistic with $k-1$ degrees of freedom. This is a measure of how well all of the independent variables affect the outcome or dependent variable. The statistic is given by G^2 =(-2 log likelihood of the null model) - (-2 log likelihood of the given model). If the $p-value$ for the overall model fit is less than the conventional 0.05, the null hypothesis is rejected and, we conclude that there is evidence that at least one of the predictor variables contributes to the prediction of the outcome (Hosmer and Lemeshow, 2000).

Wald Test

The Wald test can be used to assess the contribution of individual predictors or the significance of individual coefficients in a given model. The Wald test statistics is the ratio of the square of the estimated regression coefficient to the square of the standard error of the estimated coefficients. The Wald statistic is asymptotically distributed as a Chi-square and each Wald statistic is compared with a Chi-square with 1 degree of freedom. It is given by: $W = \frac{(\hat{\beta})^2}{(se(\hat{\beta}))^2}$, where $(\hat{\beta})^2$ is the square of the regression coefficient and $(se(\hat{\beta}))^2$ is the square of the estimated standard error of the estimated coefficient. Note that the likelihood ratio test of individual parameters is a superior criterion to the alternative Wald test when considering which variables to drop from the logistic regression model and for small sample sizes (Agresti, 2007).

Testing random effects

In multilevel logistic regression models, the significance of random effects is also tested. Mixture of chi square is used in testing random effects variance components in multilevel logistic regression models. The value of the variance under the null hypothesis lies on the boundary of the space of all possible variances (0,∞). Since parameter space is constrained to be non-negative, a one-sided test is desirable for hypothesis testing which is similar to testing that some variance components equal to zero. Constrained one-sided tests require that conventional methods for hypothesis testing be modified. Given that these are constrained one-sided tests, the distribution of the test statistic under the null hypothesis is a mixture of the X_0^2 (with all the probability mass at zero) and X^2 with corresponding df on H_a, with each of the two components of the mixture having an equal probability of 0.5.

$$X^2cal = 2 * (loglik(underHa) - loglik(underHo)) \tag{7}$$

$$p - value = \frac{1}{2} * (P(X^2(dfunderHo) > X^2calc) + P(X^2(dfunderHa) > X^2calc))$$

The correct p-value can, therefore, be obtained by simply dividing a 'naive' p-value based on X^2 with corresponding df by 2 (Austin et al., 2018).

3.3.5 Model Diagnostics

Regression diagnostics are essential to check the validity of all aspects of a regression model. Model diagnostics procedures involve both graphical methods and formal statistical tests that allow to explore whether the assumptions of the regression model are valid and decide whether we can trust subsequent inference results.

Measures of influential points

A data point is influential if it unduly influences any parts of a regression analysis, such as the predicted response, the estimated slope coefficient, or the hypothesis test results. Influential measures can be used to identify cases that are highly influential on the logistic regression estimates. Influential points affect the statistical significance as well as the strength and direction of the association between a response variable and predictor variables. The two common measures of the influence of an observation are Cook's distance and DFBETAS.

Cook's distance

It is a measure of the influence of an observation on all fitted values. The Cook's distance statistic is used to asses the influence of a data point on all predicted values. It is a measure of the change in the regression coefficient if an observation is deleted from the model. A large Cook's distance indicates that excluding a case from computation of the regression statistics changes the coefficients substantially and particular Cook's distance less than unity shows that an observation has no overall impact on the estimated vector of regression coefficients.

DFBETA(S)

This is a measure of influence of an observation on a particular regression coefficient. DFBETAS are the scaled measures of the change in each parameter calculated by deleting each observation, one at a time. This is the standardized difference in the parameter estimated due to deleting an observation, and it can be used to assess the effect of an individual observation on each estimated parameter of the fitted model. These measures are useful for detecting observations that cause instability in the selected coefficients. Values of DFBETAS less than unity imply no specific impacts of an observation on the coefficient of a particular predictor variable, while DFBETA of a case greater than 1 implies the observation is an outlier (Cook and Weisberg, 1982).

Outliers detection

Outliers generally, are extreme observations which are far away from the rest of the observations in a given data set. Detecting outliers is a common practice and it is important because outliers can affect the regression model in two ways: outliers may almost uniquely determine regression coefficients; they may also cause the standard errors of regression coefficients to be much smaller than they would be if the observation were excluded. There are two types of outliers. Outliers in the response variable represent model failure whereas outliers with respect to the predictors are called leverage points (Vittiinghoff, 2005). Data point which exercise considerable influence on the fitted model are called leverage points and leverage measures the extent to which the fitted regression model is attracted by the given data point. Outliers in x can be identified since they will have large leverage values which indicate the one case is distant from the center of all x observations and leverage considered large if it is bigger than twice the mean leverage values. An outlier may be a data point whose response y does not follow the general trend of the rest of the data and deviance residuals are often used to identify outlying value in y (large magnitude implies extreme values) and magnitude of deviance residual that is 3 or more potentially indicates an outlier.

4 Results

4.1 Descriptive Results

In this study, a total of 3,381 (weighted) teenagers (aged 15-19) were included. The mean, median and standard deviation of the ages of the respondents were 16.89, 17 and 1.34, respectively. Out of the 3,381 teenagers, 422 (12.48%) were fertile (either gave birth or were pregnant). Of those teenagers who have children, 5.75% were pregnant. About 23.81% of the teenagers in the sample lived in urban and 76.19% to rural areas. More than half of the teenagers had no exposure to mass media. The percentage of the respondents was highest for those aged 18 and lowest for those aged 19. The majority (74.24%) of the respondents had at most primary education while the rest (25.76%) had at least secondary education.

Chi-square tests of association were implemented to check the existence of an association between the outcome and predictor variables. The test results showed that all predictor variables have significant association with the variable of interest at 25% level of significance. The 25% level of significance is a screening criterion for initial variable selection based on the work of Mickey and Greenland (Hosmer and Lemeshow 2000).

Teenagers aged 18 years contributed highest and those aged 15 contributed the least fertility among all teenagers aged 15-19 years. Fertility for each category of the social and demographic characteristics have been counted and compared. Teenage fertility is higher in rural areas 382(14.83%), among uneducated teenagers 131(27.93%), and those who had no media exposure 268(16.8%). Teenage fertility was higher in Afar 23.34% followed by Somalia and Benishangul-Gumuz with 19.1% and 14.3%, respectively. Teenage fertility is also higher for teenager with poor wealth status (20.39%) and lower for teenager with poor wealth status(6.8%). The detailed distribution of fertility by social and demographic characteristics is provided in Table 2:

Table 2: Distribution of teenagers fertility status by socio-demographic characteristics.

Factor	Categories	Infertile		Fertile		Total
		count	Percent	count	Percent	
Residence	Urban	765	95.03	40	4.97	805
	Rural	2194	85.17	382	14.83	2576
Education	No education	338	72.07	131	27.93	469
	Primary	1887	87.89	260	12.11	2147
	Secondary and Above	734	95.94	31	4.06	765
Media Exposure	Not at all	1328	83.2	268	16.8	1596
	Have exposure	1631	91.37	154	8.63	1785
Wealth Status	Poor	824	79.61	211	20.39	1035
	Middle	543	85.1	95	14.9	638
	Rich	1592	93.2	116	6.8	1708
Current contra_use	No	2802	89.6	325	10.4	3127
	Yes	157	61.8	97	38.2	254
Ever-been Married	No	2627	99.43	15	0.57	2642
	Yes	332	44.93	407	55.07	739
Currently working	No	2232	87.26	326	12.74	2558
	Yes	727	88.34	96	11.66	823
Region	Tigray	243	88.04	33	11.96	276
	Afar	23	76.66	7	23.34	30
	Amhara	703	91.7	64	8.3	767
	Oromia	1025	83	210	17	1235
	Somalia	85	80.9	20	19.1	105
	Benishangul-Gumuz	29	85.3	5	14.7	34
	SNNP	608	89.3	73	10.7	681
	Gambela	8	88.89	1	11.11	9
	Harari	6	85.71	1	14.29	7
	Addis Ababa	211	97.23	6	2.77	217
	DireDewa	18	90	2	10	20
Religion	Orthodox	1316	92.22	109	7.78	1427
	Catholic	22	99.65	1	0.35	23
	Protestant	759	89.61	88	10.39	847
	Muslim	846	79.51	218	20.49	1064
	Others	15	68.18	7	31.82	22

Trend Analysis

For the trend analysis, we used the four EDHS data from 2000 - 2016. Figure 2 shows the fertility trend of all teenagers from 2000 to 2016. As presented in Figure 2, teenage fertility in 2005 and 2016 has increased (slightly) from that in 2000 and 2011, respectively. Conversely, it has decreased in 2011 as compared with that in 2005. Considering place

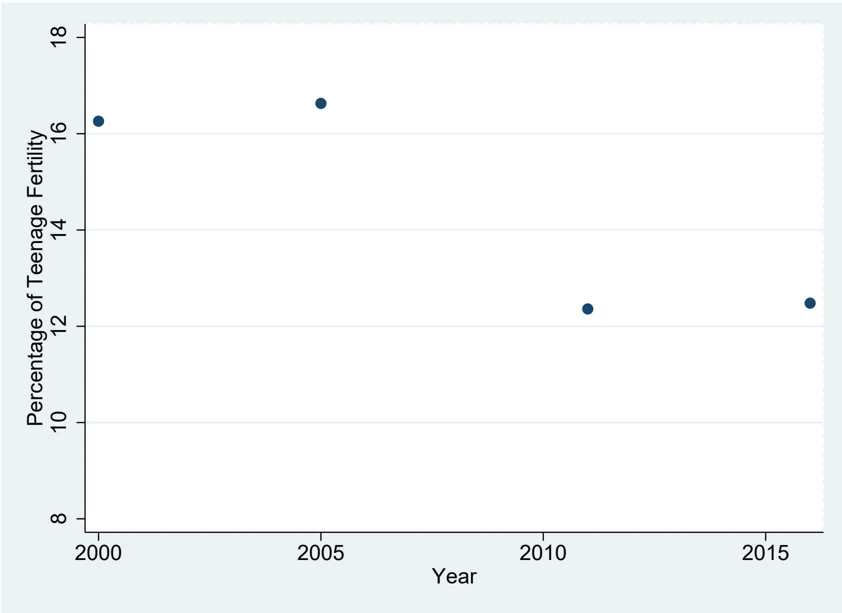

Figure 1: Teenage Fertility Trend

of residence, Figure 3 shows that teenage fertility for rural areas was consistently higher than that for urban areas between 2000 and 2016.

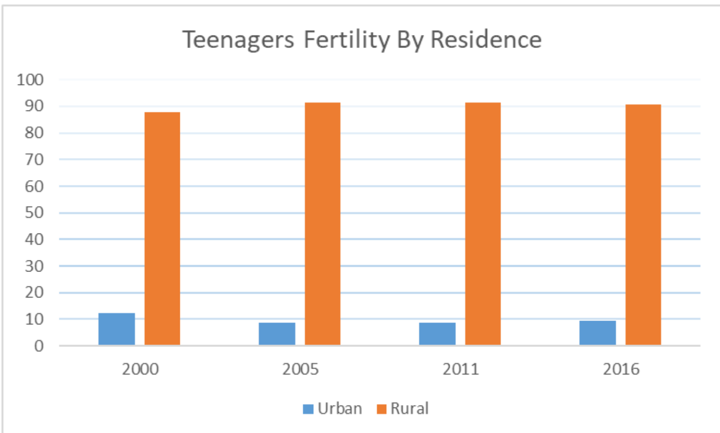

Figure 2: Teenage fertility trend by residence

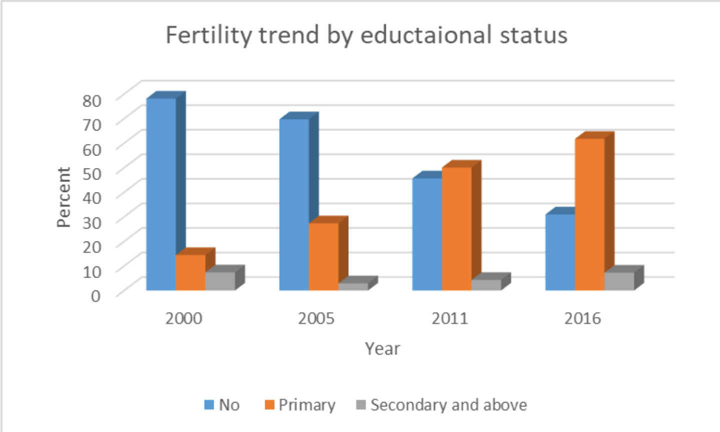

Figure 3: Teenage fertility trend by educational level

As can be seen in Figure 3, teenage fertility has shown a decreasing trend in urban areas. In rural areas, on the other hand, it had increased from 2000 to 2005 and remained the same from 2005 to 2016. Teenage fertility among those attending primary school had an increasing trend while that among uneducated teenagers had a decreasing trend during 2000 to 2016 (Figure 4).

4.2 Analytical Results

Results of Multilevel Logistic Regression Analysis

A three-level structure (with women as first level unit, household as second level unit and enumeration area as third level unit) has been used. The chi-square test was applied to assess heterogeneity between enumeration areas. The test results are $\chi^2 = 1.0e + 03$ with d.f = 635 (p=0.000). Thus, there is evidence of heterogeneity with respect to the EA's. STATA Software was used for the analysis and the results are presented below.

Table 3: Estimates of intercept-only model

Fertility status	Coefficient	Std.error	P-value	95%CI
Constant	-2.638	0.302	0.000	[-3.23, -2.045]
EA: var(cons)	1.609	0.423		[0.962, 2.694]
HH: var(cons)	1.162	0.913		[0.249, 5.419]
ICC EA	0.266	0.035		[0.203, 0.339]
ICC HH	0.457	0.113		[0.257, 0.672]

Testing the significance of random effects

The task of testing whether certain random effects should be included, which is equivalent to setting the random effect variance equal to 0, is often of interest and Wald test statistics is used. We have performed three hypotheses tests at 5% level of significance.
Testing for the significance of both random effects

- Ho: the variance of the random effects is equal to zero.
- Ha: at least one of the random effect variance is greater than zero.
  For that p-value = 0.000 $<$ 0.05 leading to rejection of the null hypothesis and we conclude that at least one of the random effects is significant

Testing for the significance of EA random effect

- Ho: the variance of the random effects due to cluster is equal to zero.
- Ha: the variance of the random effects due to cluster is greater than zero.
  For that p-value =0.000 $<$ 0.05 leading to rejection of the null hypothesis and we conclude that the random effects due to cluster is significant

Testing for the significance of household random effect

- Ho: the variance of the random effects due to household is equal to zero.

- Ha: the variance of the random effects due to household is greater than zero.
  In this case the p-value = 0.000 is very small leading to rejection of null hypothesis and we conclude that the variance of the random effects due to household is significant

As can be seen in Table 3, ICC is used to measure the proportion of variance of teenage fertility between EA and HH. The intra-class correlation coefficient for intercept only model is 0.266 and 0.457 for EA and HH, respectively. About 26.6% and 45.7% of the variation in teenage of fertility is due to variation across EA and HH, respectively.

Table 4: Estimates of the Random Intercept and Fixed Slope Model

Factor	Category	Coefficient	Std.err	P-value	95% CI
Constant		-6.953	1.199	0.000	[-9.30, -4.601]
Residence	Urban				
	Rural	-0.03	0.344	0.93	[-0.704, 0.644]
Education	No education				
	Primary	-0.084	0.279	0.763	[-0.632, 0.463]
	Secondary and above	-0.239	0.374	0.523	[-0.901, 0.494]
Media Exposure	Not at all				
	Have exposure	-0.047	0.232	0.838	[-0.502, 0.408]
Ever-been married	No				
	Yes	6.636	1.068	0.000	[4.541, 8.73]
Current contra_use	No				
	Yes	0.095	0.324	0.767	[-0.539, 0.731]
Religion	Orthodox				
	Catholic	-1.597	1.609	0.321	[-4.75, 1.556]
	Protestant	1.764	0.42	0.000	[0.941, 2.588]
	Muslim	1.288	0.334	0.000	[0.553, 1.864]
	Others	2.229	0.933	0.017	[0.399, 4.058]
Wealth status	Poor				
	Middle	-.116705	0.5059	0.818	[1.418, 7.089]
	Rich	-1.234671	0.5884	0.036	[-0.947, 1.251]
Currently working	No				
	Yes	-0.116	0.252	0.646	[-0.609, 0.378]
EA var(_cons)		0.437	0.394		[0.075, 2.562]
HH: var(cons)		2.204	1.928		[0.397, 12.245]

Model	LL	Deviance	AIC
Null	-1297.836	2595.672	2601.672
Random intercept	-613.746	1227.492	1259.492
Random coefficient	-608.416	1216.832	1250.831

The deviance of the random intercept-only model, 2595.672 reduced to 1227.492 when we include the covariates implying that the random intercept with fixed slope model is better than the random intercept only model. The AIC values also ensure this as the smaller the values of AIC the better the model is.

The deviance of the random intercept model, 1227.492 is reduced to 1216.832 implying that the random coefficient model is better than the random intercept model. The AIC values also ensure this as the smaller the values of AIC, the better the model is.

Table 5: Estimates of the Random Coefficient Model

Factor	Category	Coefficient	Std.err	P-value	95% CI
Constant		-7.63	1.34	0.000	[-10.26, -5.003]
Residence	Urban				
	Rural	0.038	0.388	0.922	[-0.723, 0.798]
Education	No education				
	Primary	-0.066	0.317	0.835	[-0.687, 0.555]
	Secondary and above	0.008	0.432	0.985	[-0.839, 0.855]
Media Exposure	Not at all				
	Have exposure	-0.034	0.261	0.896	[-0.547, 0.479]
Ever-been married	No				
	Yes	7.205	1.177	0.000	[4.897, 9.513]
Current contra_use	No				
	Yes	-0.16	0.529	0.762	[-1.96, 0.876]
Religion	Orthodox				
	Catholic	-2.951	1.759	0.093	[-6.398, 0.496]
	Protestant	2.033	0.498	0.000	[1.056, 3.01]
	Muslim	1.303	0.375	0.001	[0.569, 2.037]
	Others	2.229	0.933	0.017	[0.399, 4.058]
Wealth status	Poor				
	Middle	-.116705	0.5059	0.818	[1.418, 7.089]
	Rich	-1.196	0.632	0.04	[-2.388, -0.081]
Currently working	No				
	Yes	-0.089	0.287	0.757	[-0.652, 0.474]
EA var(_cons)		0.448	0.44		[0.064, 3.131]
var(currentlycontrause)		18.239	11.858		[5.1, 65.22]
HH: var(cons)		2.808	2.063		[0.667, 11.847]

Goodness of fit of the fitted model

Overall goodness of fit of the model was checked using the likelihood ratio test. The significant of LTR= 1368.18 , df= 7 Since the p-value of the Likelihood-ratio test is $\leq$ 0.0001. We reject the null hypothesis and conclude that at least one of independent variables is significantly related with the response variable. Significance of individual coefficients was also tested using Wald test and ever-been married, religion and wealth status were found statistically significant predictor variables at 5% level of significance.

Model Diagnostics

Model diagnostics for the random slope model was done. The results from Cook's distance and dfbetas suggest that there was no influential observation that impacts any part of the regression analysis, as the values of Cook's distance and dfbetas are below unity. Furthermore, the value of variance inflation factor indicates that there was no multicollinearity among the independent variables.

Interpretations of the results of multilevel logistic regression

The multilevel logistic regression model fit results in Table 5 reveal that ever-been married, wealth status and religion are statistically significant predictors of teenagers fertility at 5% level of significance. The odds of fertility for never married teenager is significantly differ from ever-been married teenager. The odds of fertility higher for ever-married teenagers than never-married teenagers. Table 5 also indicate that religion is significantly associated with teenage fertility. The odds of fertility for Protestant, Muslim and other religion follower teenagers are 7.637, 3.68 and 9.29 times higher than that for Orthodox teenagers. The odds of fertility for rich teenager is significantly differ from poor teenager. The odds of fertility decreased by a factor of 0.302 for rich teenagers compared to that of poor teenagers.

5 Discussion

In this study, married teenagers were more likely to be fertile than never married teenagers while controlling for others variables. This is in agreement with the findings of a study conducted by Kassa et al. (2018) that adolescents that had been married were more than twenty times more likely to start childbearing during adolescence age than adolescents who were never married. It is also consistent with the findings of the study conducted by Birhanu et al. (2019) which revealed that early married women had a higher likelihood of experiencing teenage pregnancy when compared with women who were not married before the legal age of 18. WHO(2008) also reported that about 90% of adolescent births in developing countries among 15-19 year olds occur within marriage. Even though early marriage practice is an illegal action in our country, it is still taken as normal in some cultures and in some rural areas.

This study also revealed that rich teenagers were less likely to be fertile than poor teenagers which is consistent with the finding of the study by Blum et.al (2015) which states that the poorest, least educated girls in all countries are far more likely to become adolescent mothers than their better-off peers. This disparity robs poor girls of their rights. Another study conducted by Poudel et al. (2018) suggested that adolescent pregnancy was significantly higher among woman with middle household wealth index or poor household wealth index. If teenagers were from rich household, they will have better access to education and also to receiving quality care from skilled health professionals making teenage childbearing less common among the wealthiest households. Poor teenagers might be vulnerable to sexual abuse or violence due to economic reasons. They may also engage in sex with older rich persons for economic gains.

A study by Mathewos. et al. (2018) reported that marital status and religion did not show significant association with teenage pregnancy. This finding is inconsistent with our findings.

6 Conclusion and Recommendations

6.1 Conclusion

The main objective of this study was to determine factors associated with teenage fertility in Ethiopia. The study was based on secondary data obtained from the Central Statistical Agency and the descriptive results show that about 12.48% of the sampled teenagers(weighted) were fertile.

The multilevel modeling approach enabled us to identify that teenage fertility and ever-been married, religion and wealth status are significant predictors of teenage fertility in Ethiopia.

6.2 Recommendations

Based on the findings of our study, we recommend taking the following intervention may be effective to address the problem.

- Special attention should be given to combating early marriage that will help to increase awareness about the effects and negative consequences of early marriage or having child early which in turn reduce the teenage fertility in Ethiopia.

7 References

Agresti, A. (2007), *An Introduction to Categorical Data Analysis* . John Wiley And Sons, Inc, New York.

Alemayehu .A Haider, J. and Habte, D., 2010. Determinants of adolescent fertility in Ethiopia. *Ethiopian Journal of Health Development*, 24(1).

Birhanu, B.E., Kebede, D.L., Kahsay, A.B. and Belachew, A.B., 2019. Predictors of teenage pregnancy in Ethiopia: a multilevel analysis. *BMC public health*, 19(1), p.601.

Blum, R.W. and Gates Sr, W.H., 2015. *Girlhood not motherhood.* Preventing adolescent pregnancy.

Carle A.C. Fitting multilevel models in complex survey data with design weights. *BMC Medical Research Methodology*,2009, 9:49.

Cook, R. and Weisberg, S. (1982). *Residuals and Influence in Regression* . Chapman and Hall, New York.

Central Statistical Agency (CSA). Ethiopia Demographic and Health Survey. Addis Ababa, Ethiopia, and Rockville, Maryland, USA: CSA and ICF; 2016.

Frank O. *The Demography Of Fertility And Infertility.* Geneva: World Health Organization 2017.

Hosmer, W.D. and S. Lemeshow (2000), *Applied Logistic Regression.* 2nd Ed., John Wiley and Sons, New York.

Kassa, G.M., Arowojolu, A.O., Odukogbe, A.A. and Yalew, A.W., 2018. Prevalence and determinants of adolescent pregnancy in Africa: a systematic review and Meta-analysis. *Reproductive Health*, 15(1), p.195.

Macro International Inc. measure DHS STAT compiler. 2008

Mathewos, S. and Mekuria, A., 2018. Teenage pregnancy and its associated factors among school adolescents of Arba Minch town, southern Ethiopia. *Ethiopian Journal of Health Sciences,* 28(3), pp.287-298.

National Research Council, 1993. *Social dynamics of adolescent fertility in sub-Saharan Africa.* National Academies Press.

Austin,P.C and Leckie George (2018). The effect of number of clusters and cluster size on statistical power and Type I error rates when testing random effects variance components in multilevel linear and logistic regression models. *Journal of Statistical Computation and Simulation*, 88:16,3151-3163.

Poudel S, Upadhaya N, Khatri RB, Ghimire PR (2018) Trends and factors associated with pregnancies among adolescent women in Nepal: Pooled analysis of Nepal Demographic and Health Surveys (2006, 2011 and 2016). *PLoS ONE*, 13(8): e0202107.

Population Reference Bureau(PBR). *Trends in Adolescent Fertility a Mixed Picture*,Washington, DC: May 2013.

Rai, R.K., Singh, P.K. and Singh, L., 2012. Utilization of maternal health care services among married adolescent women: insights from the Nigeria Demographic and Health Survey, 2008. *Women's Health Issues*, 22(4), pp.e407-e414.

Swartz, A., Colvin, C. and Harrison, A., 2018. The problem or the solution? Early fertility and parenthood in the transition to adulthood in Khayelitsha, South Africa. *Reproductive health matters*, 26(54), pp.145-154.

Vittinghoff, E., Glidden, D.V., Shiboski, S.C. and McCulloch, C.E., 2011. *Regression methods in biostatistics: linear, logistic, survival, and repeated measures models*. Springer Science and Business Media.

World Health Organization(WHO). Adolescent pregnancy. Department of Making Pregnancy Safer.Geneva, Switzerland:2008

World Health Organization(WHO). Early marriages, adolescent and young pregnancies.2012

World Health Organization(WHO).Adolescent Pregnancy:Issues in Adolescent Health and Development,Geneva:2004.

World health statistics overview 2019: monitoring health for the SDGs, sustainable development goals. Geneva: World Health Organization; 2019

United Nations(UN). Department of Economic and Social Affairs, Population Division. Adolescent Fertility since the International Conference on Population and Development (ICPD).Cairo:2013.

United Nations(UN).The Sustainable Development Goals Report. New York:2019.

Yussif, A.S., Lassey, A., Ganyaglo, G.Y.K., Kantelhardt, E.J. and Kielstein, H., 2017. The long-term effects of adolescent pregnancies in a community in Northern Ghana on subsequent pregnancies and births of the young mothers. *Reproductive health*, 14(1), p.178.

Appendix

Table A Chi-square test result of association between each explanatory variables and the response variable

Factor	Chi square	p-value
Residence	83.6801	0.000
Education	135.3450	0.000
Media Exposure	65.7717	0.000
Wealth Status	136.3644	0.000
Current contraceptive use	162.4660	0.000
Everbeen Married	1.6e+03	0.000
Currently working	4.4535	0.035
Religion	84.3106	0.000

Table B Results of diagnostic checking for DFBETA

Factor	Category	Minimum	Maximum
Everbeen married	Yes	-.0869883	.0757705
Religion	Catholic	-.4207258	.483022
	Protestant	-.0813315	.1383717
	Muslim	-.0959709	.0912076
	Others	-.5926518	.3783171
Wealth status	Middle	-.1005704	.1584224
	Rich	-.1213911	.0685509

Table C Testing multi-collinearity

Factor	VIF	$\frac{1}{VIF}$
Wealth Status	1.10	0.908152
Everbeen Married	1.07	0.937780
Religion	1.05	0.951734

Since the variance inflation factor is less than five there is no multicollinearity among the independent variables.

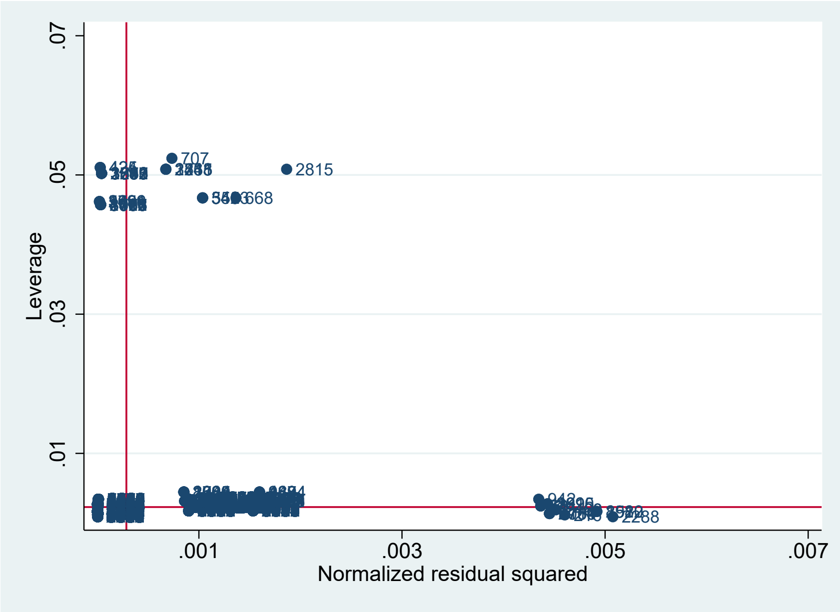

Figure 4: A graph of leverage against the (normalized) residuals squared

Points above the horizontal line have higher-than-average leverage. This plot suggests that there were observations with high leverage points. However, those observations are not influential.