52

ADDIS ABABA UNIVERSITY

OFFICE OF GRADUATE STUDIES

COLLEGE OF NATURAL SCIENCES

DEPARTMENT OF STATISTICS

# Trends and Differentials of Teenage Birth in Ethiopia

Birhanu Worku

A Thesis Submitted to

The Department of Statistics

Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics

Addis Ababa University

Addis Ababa, Ethiopia

June 2013

ADDIS ABABA UNIVERSITY

OFFICE OF GRADUATE STUDIES

Department of Statistics

This is to certify that the thesis prepared by Birhanu Worku, entitled : Trends and Differentials of Teenage Birth in Ethiopia and submitted in Partial fulfillment of the requirements for the degree of 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:

Advisor Prof. Eshetu Wencheko Signature ______ Date 25/06/2013

Examiner M. K. Sharma Signature ______ Date 25/06/13

Examiner Mekonnen Tadesse Signature ______ Date 25/06/13

---

Chair of Department or Graduate Program Coordinator

Abstract

Trends and Differentials of Teenage Birth in Ethiopia

Birhanu Worku

Addis Ababa University, 2013

Globally, each year around 16 million girls aged 15-19 give birth, accounting for around 11 percent of all births. The main objective of this study was to identify predictors of teenage birth and examine the trend of teenage birth based on data from the three EDH surveys (2000, 2005, and 2011 EDHS). Discrete-time hazard modeling was used to estimate the hazard of first birth before age 20 after controlling the effects of socio-economic factors. The results suggested that the overall likelihood of first birth before age 20 among Ethiopian women increased slightly over time in the three EDH surveys. At the individual level, women's education, especially secondary and higher educational level, had the strongest effect to delay first birth during adolescence in the three surveys. Residing in urban areas was also inversely associated with teenage birth. Having media exposure had a significant delaying influence, but the effect was low in the 2011 EDHS data. These findings reinforce our understanding that the government should continue its efforts to promote female education, especially higher education.

## ACKNOWLEDGEMENT

It gives me great pleasure to express my deepest esteem and sincere thanks to my advisor, Professor Eshetu Wencheko for his excellent scientific guidance, monitoring and constructive criticism of this thesis work.

## ACRONYMS

| | |
|---|---|
| AIC | Akaike Information Criteria |
| BIC | Bayesian Information Criteria |
| CSA | Central Statistical Agency |
| DF | Degree of Freedom |
| EDHS | Ethiopian Demographic Health Survey |
| HIV | Human Immunodeficiency Virus |
| HR | Hazard Ratio |
| DTSAM | Discrete-Time Survival Analysis Model |
| S.E | Standard Error |
| WHO | World Health Organization |
| UNICEF | United Nations Children's Fund |

# Tables of contents

## List of Tables

## List of Figures

# Chapter One

# 1 Introduction

## 1.1 Background

Early childbearing is recognized worldwide to have a profound impact on the well-being and reproductive health of young women, as well as the overall pace and direction of a country's development Alan Guttmacher Institute (AGI, 1998). Early childbearing can derail a young woman's educational prospects, reduce her long-term social and economic autonomy, and endanger both her health and that of her newborn.

Compared to adult mothers, adolescent mothers are more likely to experience maternal mortality, anemia, and obstetric complications. In addition, their infants are at higher risk for preterm birth, low birth weight, poor nutritional status and fetal death. However, these effects should be viewed within the context of maternal and child health conditions in a given society. In poor countries, the health of women and children is also influenced by a range of social and economic factors such as the mother's education, access to health care services, decision-making power, acceptance of contraceptives, and employment opportunities (Gill et al 2007; Taffa and Obare 2004).

Early childbearing often results from child marriage. Almost all adolescent birth occurs within marriage in Asian and North African countries, as do around 70-80 percent in Sub-Sahara Africa and Latin America and the Caribbean (WHO, 2008). Marriage increases the frequency of exposure to unprotected sex for young girls, which increases their risks of becoming pregnant or contracting sexually transmitted diseases and cervical cancer (Nour, 2006). According to studies in Kenya, Zambia and Zimbabwe, married adolescents are more likely to be infected with HIV than other sexually active girls of their own age (Clark 2004; Clark et al, 2006; Gavin et al. 2006). Behaviors that increase this risk include higher frequency of sex, lower likelihood of using condoms, and a greater likelihood of having older partners (Clark et al, 2006). According to UNICEF 2012 reports on adolescents, 2.2 million adolescents 10-19 years old are living with HIV globally, and 1.8 million in Sub-Sahara Africa. Moreover this report reveals that, among adolescents 15-19 years old in the developing world (excluding China), a higher percentage of girls (11 percent) than boy (5 percent) had sex before the age of 15. This indicates that early sex can result in early childbearing, and it increases the risk of HIV infection.

Globally, each year around 16 million girls aged 15-19 gave birth, accounting for around 11 percent of all births. Almost 95% of these births occur in developing countries (2011). They range from about 2% in China to 18% in Latin America and the Caribbean. Half of all adolescent births occur in just seven countries: Bangladesh, Brazil, the Democratic Republic of Congo, Ethiopia, Nigeria, India and the United States (WHO 2008).

Countries of Latin America and the Caribbean and Sub-Saharan Africa have the highest proportion of adolescent births (UNICEF, 2012 report). Approximately 95 percent of adolescent births occurred in low and middle-income countries (United Nations Department of Economic and Social Affairs, 2011). Bangladesh, India, and Nigeria alone account for one in every three of the world's adolescent births and the only industrialized country among the top 10 countries with the highest number of adolescent births is the United States (UNICEF, 2012).

Having this background information and, because of its far-reaching individual and social consequences, adolescent childbearing remains a major concern for developed and developing nations alike. In a country where marriage is universal and occurs at young ages, early teenage childbearing is of prime public health importance. Accordingly, it is relevant to assess the trends and characteristics of teenage birth using three nationally representative datasets available based on the three 2000, 2005, and 2011 Ethiopian Demographic and Health Survey. Such a study can inform policy related to maternal health issues, with special emphasis on teenage women.

## 1.2 Statement of the problem

Adolescent girls who give birth each year have a much higher risk of dying from maternal causes compared to women in their 20s and 30s. These risks increase greatly as maternal age decreases, with adolescents under 16 facing four times the risk of maternal death as women over 20. Moreover, babies born to adolescents also face a significantly higher risk of death compared to babies born to older women (WHO, 2008). Additionally, WHO revealed that adolescent childbearing has a negative impact on these three dimensions: health of the adolescents and their infants; individual social and economic effects; and societal level impacts.

Teenage childbearing also negatively impacts the survival of newborns. This reinforces that the age of the mother at the time of the first birth is an important factor for infant and child survival. For instance, Mondal *et al.* (2009), found that the most significant predictors of neonatal, postneonatal, and child mortality levels were mother‘s age at first birth along with other covariates (immunization, ever breastfeeding, and birth interval). Additionally, studies have shown that there is a relationship between age at first birth and infant and child mortality (Bicego and Boerma, 1993). Results of such studies indicate that the younger the age of the mother at the birth of the first child, the higher the chances of the child dying due to complications during childbirth. For example, in Ethiopia, infant mortality was higher for births to mothers under age 20 (EDHS 2011 report). According to the data from 2000 EDHS, 37.8% women of aged 20-24 at time of interview had their first birth from age 15 to 19. Similarly, in the 2005 EDHS, 39.4%, and in the 2011 EDHS, 38.5% of women aged 20-24 had first birth from age 15 to 19. This indicated that despite the downward trend from 2005 to 2011, teenage birth remains very prevalent in Ethiopia. Given this as a backdrop, this study will focus on the trends and differentials of teenage birth using 2000, 2005, and 2011 Ethiopian Demographic Health Survey.

Specifically, this study will focus on the following:

1. What is the trend of having first birth before age 20 in Ethiopia and what characteristics were associated with those women who had first birth before age 20?
2. Additionally, in this study, we are not only interested in 'whether' a woman had first birth before age 20, but also in 'when' they had first birth (age at first birth).

### 1.3 Objectives

The general objective of the study is to analyze trend and differentials of teenage birth in Ethiopia.

The specific objectives of the study are:

- To study the trends of teenage childbearing
- To identify socioeconomic and demographic factors determining teenage birth

### 1.4 Significance of the study

This study is meant to contribute to understanding African teenage childbearing by examining the situation prevailing in one particular country, Ethiopia. It is important to understand when childbearing begins, and factors associated with teenage birth so that the impact of the National Population Policy can be assessed. Also, the analysis of new data could be used to give a better understanding of teenage childbearing trends in Ethiopia in the recent past.

### 1.4 Limitation of the study

- The study heavily depends on information pertaining to the timing of event. That is age at first birth. On the other hand, as with all observational studies, responses in the DHS surveys are not immune to errors such as recall errors due to memory lapses and event omission (both deliberate and accidental). Therefore, readers are advised to cautiously interpret the results of this study.

- Several studies have shown that women with early menarche have earlier first births. This study did not consider the variable age at menarche, as the three EDHS data did not provide any information on it.

# CHAPTER TWO

## LITERATURE REVIEW

Factors influencing fertility can be classified into two groups: (1) proximate variables and (2) "background" variables (Bongaarts, 1982). While the proximate variables consist of biological and behavioral factors, which have direct influence on fertility, the background variables comprised of socioeconomic and environmental factors, which affect fertility only indirectly by modifying the proximate variables. Although Davis and Blake (1956) were the first to identify a set of 11 proximate determinants, Bongaarts (1978) reclassified this list into eight variables – proportions married, contraception, induced abortion, lactational infecundability, frequency of intercourse, sterility, spontaneous intrauterine mortality, and duration of fertile period.

Because lactational infecundability only applies to women who have a previous child, it is not appropriate to consider it for the analysis of age at first birth. Given the lack of consistency and reliability of data, proximate variables like, sterility, spontaneous and induced abortions are also not considered in this study. The three EDHS of 2000, 2005, and 2011 data do not provide full information about age at menarche and practice of contraception in order to delay age at first birth. Although age at first marriage is often used as a proxy for first exposure to sexual intercourse, the two events do not necessarily coincide. But Ethiopian women generally begin sexual intercourse at the time of their first marriage (2011Ethiopian DHS final report). Therefore the present paper is focused on the effect of background variables. Thus, the following review is limited to individual, economic, and cultural factors, which are known to affect teenage fertility.

One of the most consistent findings of analyses of fertility behavior in developing countries is a strong correlation between the level of women's education and fertility behavior. Schooling of women, which is often considered an indicator of socio-economic development, is found to be associated with delayed marriage, increased contraceptive use, decreased family size, and reduced infant mortality. Yet, the effect of schooling on adolescent childbearing is not straightforward. Even though there is evidence that higher levels of education are associated with lower probability of giving birth during adolescence (Islam, 1999) and elsewhere (Gupta and Leite, 1999), teenage girls may discontinue schooling after getting married and/or getting pregnant. Similarly, they can postpone marriage and delay childbearing in order to complete schooling.

Singh et al (2001) studied the socioeconomic disadvantages and adolescent women's sexual and reproductive behavior, in the case of five developed countries. According to their study, adolescent childbearing was more likely among women with low income and low level of education than among their better-off peers. Moreover, the study showed that young women who had little education were more likely to initiate intercourse during adolescence than those who were better educated.

Singh (1998) studied the levels and trends of the rate of adolescent childbearing, the timing of first birth, and births to unmarried women for 43 developing countries. According to the results of the study, higher education was associated with lower rates of adolescent childbearing, but others socioeconomic changes cancel or reduce this effect in several countries.

A study by Tim et al (2002) showed that first intercourse, first union and first birth of Latin American women with secondary level of schooling were less likely to experience early marriage or parenthood relative to those with no schooling.

Olausson et al (2001) studied teenage childbearing and long-term socioeconomic consequences in Sweden. Using multilevel logistic regression techniques, the result of the study showed that teenage motherhood was positively associated with low educational attainment.

Nahar and Min (2008) examined the trends and determinants of adolescent childbearing in Bangladesh using four sets of Bangladesh Demographic and Health Survey data collected during 1993/94, 1996/97, 1999/00, and 2004. The study used discrete-time multilevel hazards model to estimate the hazard of first birth before age 20 after controlling the effects of other individual and household factors. The results showed that women's education, especially higher education has the strongest effect in delaying first birth during adolescence.

Gupta and Leite (1999) used data from three demographic and health survey to examine trends and determinants of fertility behavior among adolescents in Brazil. The results of the study showed that high level of a young woman's education was strongly associated with delayed childbearing in the region.

Gupta and Mahy (2003) conducted a study in eight sub-Sahara African countries based on data drawn from Demographic and Health Survey on adolescent reproductive behavior. The results of the study showed that education (Grade 8 and above) was found to have consistently and significantly reduced the risk of adolescent childbearing in all countries studied. However, a study by Cesare and Rodriguez (2006) had shown an ambiguous effect of education. In their model where only socioeconomic variables were included, education had a significant effect whereas when all significant socioeconomic variables and proximate determinants were incorporated into a model to predict adolescent childbearing education had no effect.

Kamal (2012) studied the fertility of adolescent mothers in Bangladesh using the 2007 Bangladesh Demographic and Health Survey data. The multilevel logistic regression analysis revealed that women's education had a significant depressive influence on the probability of adolescence childbearing or overall, women's secondary or higher education acted as catalyst toward delayed childbearing.

Another socio-economic factor that is thought to be associated with women's fertility behavior is women's place of residence (Schultz, 1981). As hypothesized by Singh (1998), young girls living in urban areas may have greater motivation to attain higher education and to work for wages, as well as a greater availability of work opportunities, and thus, were less likely to have teenage pregnancy compared to their rural counterparts. Limited analyses from Bangladesh support this hypothesis and suggested that early childbearing among teenagers was less common among urban residents than rural residents (NIPORT, Mitra and Associates, and ORC Macro, 2005).

A review of adolescent childbearing (Singh, 1998) done in 43 developing countries documented that the level of adolescent childbearing is lower in urban areas compared to rural areas, but there are a number of exceptions (e.g. Namibia, Sri Lanka, Trinidad, Tobago, and Turkey). This study showed that urban Botswanan woman experienced a high level of adolescent childbearing than did rural Botswanan women.

Tim et al (2002) showed that urban residence delayed entry into sexual activity, marriage and parenthood in Latin America. However, once education is controlled for, the effect of residence reverses – that is, urban residence was found to be associated with early transition ages.

Gupta and Mahy (2003) studied adolescent childbearing in eight sub-Sahara African countries. The results showed that residing in urban centers had a sizeable effect on the chance of adolescent childbearing only in three of the eight countries considered in the study. They found that the association between place of residence and risk of early childbearing was significant in Côte d'Ivoire, Ghana, and Senegal. Girls living in urban areas were over 30 percent less likely to have a first birth before age 18 as compared to rural residents.

Kamal (2012) showed that women's place of residence was found as an important determinant of adolescent motherhood in Bangladesh. One of the possible reasons was that government provided female stipend programme in Bangladesh which contributed to substantially reducing the wide gap of educational attainment of adolescents of rural and urban areas.

On the other hand, results from studies done in Brazil and Colombia indicated that place of residence had no significant effect or loses its effect when socioeconomic level was controlled (Gupta and Leite, 1999; Cesare and Rodriguez, 2006).

Chandrasekhars (2010) studied factors affecting age at marriage and age at first birth in India. Using cox proportional hazard model, the results showed that women who lived in the countryside are more likely to have married earlier and have children earlier compared with married women living in a large city.

The effect of income or economic status on fertility behavior is harder to predict than that of women's education or place of residence (Dreze and Murthi, 2001). The literature has argued that the effect of income depends on whether children are considered as an economic burden or valued as a productive asset. However, this argument may not be applicable to the timing of teenage birth as the effect of economic status may be different for the first birth compared to overall fertility status.

An economic analysis of fertility by Davis et al (1993) showed that increases in family income are associated with a few number of children. Studies done on adolescent childbearing support the claim that adolescent childbearing was associated with income and poor earning opportunities, occurring among poor and unemployed rather than to all women (Buvinic,1998; Singh *et al.,* 2001).

Mass media plays an important role in disseminating information and bringing social changes with respect to attitude towards fertility behavior. Analysis of DHS data from Northeast Brazil supported this assumption and found that access to media was the most important predictor of fertility among women age 20-30 years (Gupta and Leite, 1999). Exposure to media provides the opportunity to be acquainted with new ideas and knowledge that is useful in various aspects of everyday life (CSA Ethiopia and ORC Macro, 2006).

However, the survey conducted in eight sub-Sahara African countries by Gupta and Mahy (2003) reported that exposure to media had no significant effect on teenage childbearing in six of the eight countries considered in the survey. They found that regular radio listening habits were inversely associated with the probability of an adolescent first birth in Côte d'Ivoire and Zimbabwe. Additionally, Nahar and Min (2008) did not find a positive effect of media exposure to decrease the chance of having a teenage birth.

Differential fertility behaviors have been reported among different religious and cultural groups throughout the world. In particular, studies conducted in Brazil suggested that Catholics had higher fertility rates compared to other religious groups (Gupta and Leite, 1999). Similarly, higher fertility rates were reported among Indian Muslims compared to other religious groups (Dreze and Murthi, 2001). Yet, religious affiliation as a determinant of teenage fertility needs to be examined.

Chandrasekhars (2010) found that compared to women who reported to be Hindus, Muslims were likely to marry earlier. Christians and women from other religious groups were likely to marry later compared to Hindus.

Using evidence from 2007 Bangladesh Demographic and Health Survey data, religion did not show a net effect on childbearing among adolescents, but it appeared as a vital predictor for teenage motherhood for elder women (Kamal 2012). Non-Muslim women were significantly ($p<0.001$) 45.3% less likely to have childbirth at adolescence than their peer Muslim counterparts.

Singh (1998) reviewed the levels and trends in rate of adolescent childbearing, the timing of first birth and births to unmarried women in 43 developing countries. The review reported that a decline in adolescent childbearing was observed in many of the countries studied. The study found out that younger cohorts were more educated and lived in urban areas. Besides, the decline observed was more common among the better educated and those who lived in urban areas. However, a small increase in adolescent childbearing was observed in five sub Saharan Africa and four Latin American countries even among the better educated women.

Very little has been done to examine the effect of background factors in explaining fertility among Ethiopia women. Tewodros et al (2010) studied the socioeconomic and demographic determinants of adolescent fertility in Ethiopia using the 2005 EDHS data for women in the 15-19 age groups. The multivariate logistic regression analyses showed that education, place of residence and current age were strong determinants of adolescent fertility. Working status after controlling for selected socio-demographic variables also had an impact. Moreover the fertility rate of women in Addis Ababa and other urban area did not differ much. Those with primary or secondary level of education had lower probability of childbearing during their adolescence period than those without education.

To the best of our knowledge to date, no study had focused on the effects of background factors in explaining teenage childbearing in the three EDHS data. Therefore, the objective of this study is to study the trends as well as differentials of teenage birth using data obtained from the three EDHS of 2000, 2005, and 2011.

# CHAPTER THREE

# DATA AND METHODOLOGY

## 3.1 Data source

In this study data from the 2000, 2005, and 2011 Ethiopian Demographic and Health Survey have been used. The three surveys were conducted by the Central Statistical Agency (CSA) under the auspices of the Ministry of Health 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. The primary objectives of the EDHS were to provide up-to-date information for planning, policy formulation, monitoring, and evaluation of population and health programmes in the country.

The 2000 Ethiopia Demographic and Health Survey (DHS) was the first of its kind to be conducted in the country. In that survey a nationally representative data set was obtained through interviews with 15,367 women aged 15-49. Among these 6,428 were aged 15-24. In the 2005 EDHS a total of 1, 4070 women were interviewed including 5,869 women aged 15-24. From the third Demographic and Health Survey conducted in 2011 information on fertility and family planning was obtained. This survey included 1, 6515 women aged 15-49 of whom 6,857 were of age 15-24. Since the primary objective of this study is about fertility behavior of the youngest cohort, we focused on female respondents aged 15-24 at the time of each survey.

## 3.2 Study variables

### 3.2.1 The Response Variable

The response variable in this study is "age at first birth before age 20 in completed years". We define $y_{it}$ as the binary response at age t (t=15-19, in completed years) of woman i having a first birth. $y_{it}$ is set to 1 if the women had her first child at age t, and set zero otherwise.

### 3.2.2 Predictor Variables

Predictor variables (factors) included in the model which are assumed to determine teenage fertility were:

a. Woman's religion: Coptic orthodox, Protestant, Muslim, and "Others". The last group included Catholics and followers of traditional beliefs.
b. Frequent media exposure: measured by asking respondents whether they watched television, radio or read newspaper on a weekly basis.
c. Woman's educational attainment: recorded as no education, primary, secondary and above.
d. Place of residence: Urban and Rural
e. Occupation: refers to working status of women and or the type of job a woman was engaged in at the time of the survey. It is classified as 'Not working', 'Agricultural worker', and 'Nonagricultural worker.'
f. Region: refers to the 9 regional administrations and two city administrations of Ethiopia in which a woman was living at the time of the survey.

**Table 1: Description of Variables included in the Analysis**

Response Variable

| Variable | Representation of variable | Category |
|---|---|---|
| Time to first birth of a women before age 20 | T | Discrete-time (15,16,17,18,19 years) |

Predictor Variables

| Variables | Representation of variable | Coding |
|---|---|---|
| Woman's religion | REL | 0 = Orthodox<br>1 = Protestant<br>2 = Muslim<br>3 = Others |
| Media exposure | MED | 0 = Not at all<br>1 = Exposure |
| Woman's educational attainment | EDU | 0 = No education<br>1 = Primary<br>2 =Secondary and above |
| Place of residence | PLR | 0 = Urban<br>1 = Rural |
| Occupation | OCC | 0 = Not working<br>1 = Agricultural worker<br>2 = Nonagricultural worker |
| Region | REG | 1 = Tigray,<br>2 = Affar,<br>3 = Amhara,<br>4 = Oromiya,<br>5 = Somali,<br>6 = Benishangul-Gumuz,<br>7 = SNNP,<br>8 = Gambela, |

| | | |
|---|---|---|
| | | 9 = Harari,<br>10=Addis Ababa<br>11= Dire Dawa |

### 3.3 Methods of data analysis

### 3.3.1 Discrete-Time Hazard Analysis

Births occur in the reproductive age 15-49 years, and in some exceptional cases outside this age. Thus, for the analysis of birth histories it is appropriate to use continuous-time models such as the Cox model. However, data about birth histories are typically collected via retrospective surveys. In such surveys it is common practice to record dates in large grouped-time intervals such as months or years. The application of continuous-time models to grouped-time survival data is not recommended because of the problem of the large number of ties (i.e., more than one individual experiences an event at the same time). To overcome difficulties that continuous-time methods have with these grouped-time data, alternative methods have been developed (Allison, 1982). A popular alternative is the discrete-time approach, where time is treated as though it were truly discrete (Myer, Hankey and Mantel, 1973; Brown, 1975).

Discrete time hazard models (in multi-level form) are used to estimate the probability of teenagers having a first birth. This approach allows considerable flexibility in handling time-varying covariates (in particular, a woman's age) (Allison, 1982). Another advantage of discrete-time hazard models is that they allow to fit censored observations (that is, teenagers aged 15-19 who had not yet completed adolescence at the time of the survey), as well as women aged 20-24. The model is essentially a logistic regression model with the response variable being the log odds of a women having had a first birth at age t (t=15, 16, 17, 18, 19).

In a situation described here the appropriate statistical tool is a discrete-time survival model (e.g. Allison, 1982; Singer & Willett, 1993), where the time for subject i is represented by a discrete random variable $t_i$ assuming values that are positive integers.

The discrete-time hazard probability is the conditional probability that an individual i will experience the event of interest at time t given that the individual has not experienced the event of interest in any earlier time intervals (Singer & Willett, 1993).

That is:

$$h_i(t) = P(T_i = t | T \geq T_i) \tag{1}$$

where i = the i[th] subject in any of the surveys

These conditional probabilities, the $h_i(t)$, are the fundamental parameters of the discrete-time survival process. As the central focus of all analysis, we estimate their values and investigate their dependence on selected covariates in this study. In the setting of age at first birth, $h_i(t)$ is the probability that a teenager i gave first birth in age t given that she had not given birth before time t. This function can be visualized as a plot, against time, of the event occurring in each time period.

Inference methods for survival analysis allow for right censoring. A teenager is right censored at age t if the observation period ends before experiencing the event of interest (first birth in this case). Thus the observation period for this subject is not $T_i = t$, but rather $T_i > t$. The end of the observation period may be determined by the design of the survey. In this study since EDHS is a retrospective (a single interview) survey the observation period is ended by design at the day of the interview.

With right censoring, the observation of subject $i$ is represented with an ordered pair $(t_i, y_{it})$, where $t_i$ is the time recorded and $y_{it}$ is an indicator of the occurrence of the event of interest. Thus $y_{it}$ =1 means that $t_i$ is uncensored ($T_i = t$), while $y_{it}$ = 0 means that $t_i$ is censored ( $T_i > t$ ). Censored times cannot be deleted, nor treated as uncensored times. The standard estimation methods for survival models use the censored times under the assumption of non informative censoring. Informally, censoring is non informative when, conditionally on the observed covariates, the end of the observation period does not depend on the hazard.

We next include a set of q predictors, to equation (1). These predictors allow researchers to characterize the individuals in the population. We denote the q predictors in time period t for the i$^{th}$ individual with the vector $X_{it} = (x_{1it}, x_{2it}, \ldots x_{qit})'$. The discrete-time hazard function for individual i, in time period t, with q predictors is given as follows (Singer & Willett, 1993):

$$h_i(t \mid x_{it}) = P(T_i = t \mid T_i > t, X_{1it} = x_{1it}, X_{2it} = x_{2it}, \ldots X_{qit} = x_{qit}) \qquad (2)$$

i= individual observation, t=15, 16, 17, 18, 19

where the vector $X_{it}$ includes all the covariates of subject *i* at time *t*. The covariates can be time-invariant or time-varying. Time-varying covariates are extremely useful in building a proper model for the hazard, but they are rarely available in practice because of the difficulty to measure them accurately, especially in retrospective surveys. So, only time-invariant covariates were considered in this study.

#### 3.3.1.1 A statistical model for discrete-time hazard

Although equation (2) shows that the hazard depends on each individual's value on a vector of predictors, it does not specify the functional form of the dependence. Thus, in this section a description of a formal model of a hypothesized relationship between the population hazard probabilities and predictors will be provided.

The most popular choice to specify how this hazard depends on time and the explanatory variables (Cox, 1972; Myers, Hankey, and Mantel, 1973; Byar and Mantel, 1975; Brown, 1975; Thompson, 1977; Mantel and Hankey, 1978; Allison, 1982; Singer and Willett, 1993) is the logistic regression model. The model represents the log-odds of event occurrence as a function of predictors and also has the attributed of baseline profile risk and a shift parameter that captures the effect of the predictors on the baseline profile (Singer and Willett, 1993). Therefore our proposed population discrete-time hazard model is:

$$h_i(t)=\frac{1}{1+\exp[-\{(\alpha_1 A_{1it}+\alpha_2 A_{2it}+...+\alpha_T A_{Tit})+(\beta_1 X_{1it}+\beta_2 X_{2it}+...\beta_q X_{qit})\}]} \quad (3)$$

t=15, 16, 17, 18, 19

Here $A_{1it}, A_{2it}, \ldots, A_{Tit}$ are a sequence of dummy variable, with values $(a_{1it}, a_{2it},..a_{Tit})$ indexing time period. $\alpha_1, \alpha_2, \ldots, \alpha_T$ are the intercept parameters that capture the baseline level of hazard in each time period. The slope parameters $\beta_1, \beta_2, \ldots, \beta_q$ describe the effects of the predictors on the baseline hazard function, albeit on a logistic scale (Singer and Willett,1993). T refers to the last time period observed for anyone in the sample. If $t_i$ represents the last time period when individual i was observed (and at which time she was either censored or experienced the target event, then T=sup $\{t_i\}$.

Taking the logistic transformation of the both sides of (3) we obtain

$$\ln\left(\frac{h_i(t)}{1-h_i(t)}\right)=(\alpha_1 A_{1it}+\alpha_2 A_{2it}+...+\alpha_T A_{Tit})+(\beta_1 X_{1it}+\beta_2 X_{2it}+...+\beta_q X_{qit}) \ . \quad (4)$$

t=15, 16, 17, 18, 19

This form assumes that the predictors are linearly associated with the logistic transformation of hazard (logit-hazard), not with the hazard themselves, nor with the natural logarithm of the hazard probabilities.

We also notice that the discrete-time hazard model contains no single intercept, instead the alpha parameters act as multiple intercepts, one per time period. When the values of all the covariates, $X_1, X_2, \ldots, X_q$ are set zero, the population discrete-time hazard model depends only on $\alpha_1, \alpha_2, \ldots \alpha_T$ represent the population baseline logit-hazard function because it captures the time-period by time-period conditioning log odds that individuals whose covariate values are all zero (baseline group) will experience the event in each time period, given that they have not already experienced it (Singer and Willett,1993).

#### 3.3.1.2 Estimation technique for the parameters of the discrete-time hazard model

Let $y_{it}$ be a dichotomous indicator variable of the occurrence of the event of interest, that is, whether a teenager female gave birth at time t or not. $y_{it}$ is 0 if individual i in time period t did not experience the event of interest and $y_{it}$ is 1 if individual i in time period t experienced the event of interest. There will also be instances when an individual does not experience the event of interest before the observation time ends, and those individuals must be censored. Let $C_i$ be a dichotomous indicator variable that describes if an individual was censored or not. Therefore we have, $c_i = 0$ if individual i has not been censored and $c_i = 1$ if individual i has been censored.

The maximum likelihood method is used to estimate the parameters $\alpha_1, \alpha_2, \ldots, \alpha_T$ and $\beta_1$, $\beta_2, \ldots, \beta_q$ in equations (3) and (4) and we therefore get an estimate for $h_i(t)$. The likelihood function must be constructed in two parts because of censoring. The two parts of the likelihood function deal with first the uncensored individuals, that is, the probability that the individual experienced the event of interest in time period $t_i$, and the censored individuals, that is, the probability that the individual experienced the event of interest after time period $t_i$.

That is the contribution of subject *i* to the likelihood is different if the time is uncensored or censored:

$$\text{For uncensored } (y_i = 1)\text{: } P(T_i = t_i) = \prod_{u=1}^{t_i - 1} [1 - h_i(u)] \times h_i(t_i) \tag{5}$$

$$\text{For censored } (y_i = 0): P(T_i > t_i) = \prod_{u=1}^{t_i} [1 - h_i(u)] . \tag{6}$$

Here $t_i$ represents the last time period when individual i was observed.

Now following Alison (1982), Singer and Willett (1993) we have the following:

Assuming that individuals in the sample are independent (given their $x_{1it}, x_{2it}, \ldots, x_{qit}$ values), the likelihood function is simply the product of the probabilities of observing the sample data, $P(T_i = t_i)$ in the case of the uncensored individuals( $c_i = 0$) and $P(T_i > t_i)$ in the case the censored individuals ( $c_i = 1$):

$$L = \prod_{i=1}^{n} [P\{T_i = t_i\}]^{1-c_i} [P\{T_i > t_i\}]^{c_i} \tag{7}$$

Now substituting (5) and (6) into (7), and taking logarithm we have

$$l = \sum_{i=1}^{n} \left[ (1-c_i)\ln\left(\frac{h_i(t_i)}{1-h_i(t_i)}\right) + \sum_{t=1}^{t_i} \ln(1-h_i(t)) \right] \tag{8}$$

The event-history indicator $Y_{it}$ can be used with equation (8), and we have the following equation:

$$\sum_{t=1}^{t_i} y_{it} \ln\left(\frac{h_i(t)}{1-h_i(t)}\right) = \begin{Bmatrix} \ln\left(\frac{h_i(t_i)}{1-h_i(t_i)}\right) when, c_i = 0 \\ 0, when, c_i = 1 \end{Bmatrix}$$

$$= (1-c_i)\ln\left(\frac{h_i(t_i)}{1-h_i(t_i)}\right) \quad . \tag{9}$$

Substitute (9) into the first term inside the bracket of equation (8) which eliminates the censoring indicator $c_i$ from the log-likelihood and replacing it by the dichotomous realization of the event-history process the $y_{it}$ we have the following equation.

$$l = \sum_{i=1}^{n} \left[ \sum_{t=1}^{t_i} y_{it} \ln\left(\frac{h_i(t_i)}{1-h_i(t_i)}\right) + \sum_{t=1}^{t_i} \ln(1-h_i(t)) \right].$$

This can be written as:

$$l = \sum_{i=1}^{n} \sum_{t=1}^{t_i} \left[ \ln\left(\frac{h_i(t_i)}{1-h_i(t_i)}\right)^{y_{it}} + \ln(1-h_i(t)) \right]$$

If we combine like terms and take the antilog we have:

$$L = \prod_{i=1}^{n} \prod_{t=1}^{t_i} h_i(t)^{y_{it}} \left(1 - h_{it}\right)^{(1-y_{it})} \tag{10}$$

Equation (10) is the likelihood function for the discrete-time hazard process in terms of the data (the $y_{it}$ ) and the hazard probability parameters (the $h_i(t)$ ).

As demonstrated by Allison (1982), Brown (1975), and Laird and Oliver (1981), the equivalence of the likelihood functions of the discrete-time hazard model in (10) and the independent Bernoulli trials model allows us to treat the N dichotomous observed values $y_{it}$ as a collection of independent dichotomous variables with a hypothesized logistic dependence on predictors. They can be regarded as the values of the outcome variable in a logistic regression analysis of the time-period indicators (A1 through A5) and covariates X. This provides a simple method of obtaining maximum likelihood estimates of $\alpha_1, \alpha_2, \ldots, \alpha_T$ , $\beta_1, \beta_2, \ldots, \beta_q$ and hence the $h_i(t)$ using standard logistic regression analysis software (Singer and Willet 1993). Because computer software for conducting logistic regression analysis is so widely available, we will illustrate the fitting of hazard models via standard logistic regression approach, rather than via direct maximization of the likelihood in (10).

#### 3.3.1.3 Constructing the Person-Period Data

In a typical data set, each person (case) has one record of data. Discrete-time survival analysis model (DTSAM) requires a person-period format; that is, each person may have a different number of records depending on the duration of observation. So, the first step to conduct DTSAM is to convert the data into a person-period data format. In the converted person-period data set, different cases may have a different number of records depending on how long it takes to experience the event (time to first birth). Therefore before we conduct discrete-time survival analysis we transform the standard one-person, one-record data set (the person-period data set) as shown in Table 4.4, Appendix A.

#### 3.3.1.4 Baseline Hazard Models

Prior to estimating multi-level discrete-time models, it is useful to estimate a set of preliminary models for comparison purposes. We begin by estimating a simple discrete-time hazard model using a standard logistic regression model that includes only a set of age dummy variables (A1 through A5, see Table 4.4) and no intercept model (Model 1, represented by equation (11)) as follows:

$$\psi_{it} = \ln\left(\frac{h_{it}}{1-h_{it}}\right) \tag{11}$$

$$= \sum \alpha_t (AGE_{it})$$

where $h_{it}$ is the hazard of giving first birth for person i at year t, and $AGE_{it}$ is a dummy variable for age $t$ for person $i$. The estimated coefficients, $\alpha_i$ give the shape of the baseline logit-hazard curve (Reardon et al., 2002).

#### 3.3.1.5 Discrete-time hazard model with covariates

In the next model, we add demographic covariates to model 1 in equation 11. Model 2 is represented by the following equation:

$$\psi_{it} = \sum \alpha_t (AGE_{it}) + \beta X_i \tag{12}$$

where $X_i$, is a vector of time-invariant covariates for teenager female i. Model 2 was used to estimate the effects of the demographic covariates on the logit-hazard curve (Reardon et al., 2002).

### 3.4 Model Comparisons

It is useful to be able to judge whether a model is a good fit to the data. For this study, test of goodness of fit uses the deviance. The maximum likelihood procedure produces a statistic called the deviance, which indicates how well the model fits the data. The test compares the deviance (-2 log likelihood) of two models by subtracting the smaller deviance (model with more parameters) from the larger deviance (model with lower parameters). The difference is a chi-square with the number of degrees of freedom equal to the difference in parameters in the two models. Similarly, model comparison is also examined using Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC). The smaller the value, the better of the model will be.

# CHAPTER FOUR

# RESULTS

The data analysis was done using R3.0.0 statistical (software) package. The results of the analysis are divided into two sections: descriptive analysis results, and results of discrete-time hazard model analysis.

## 4.1 Descriptive Analysis

The major demographic and socioeconomic background characteristics of the respondents with first birth before age 20 are presented in Table 4.1 in the Appendix A. As shown in the table, the proportion of women with first birth before age 20 in the three surveys increased from the 2000 survey to the 2005 survey then dropped slightly in 2011. The overall proportion of women who gave first birth before age 20 in 2000 decreased by 2.63% in 2011. During the same period, there has been significant improvement in attainment of educational level: in 2000 only 8.2% of women age 15-24 had secondary or more than secondary education and slightly increased in 2005 (8.7%) and decreased in 2011(5.7%).

There has been a major upward shift in the improvement of media exposure. In 2000 only 35.5% of women age 15-24 reported to be exposed to one of the three media (reading newspaper or magazine, listening to radio and watching television) and increased to 57.8% in 2011.

The proportion of first birth before age 20 of women varied from region to region. For example, in the 2011 survey the highest percentage of first birth before age 20 was observed in Oromiya (14.3%) followed by Amhara (12.7%) whereas the lowest percentage was recorded in Addis Ababa (2.7%) and followed by Dire Dawa (5.2%). Specifically, Benishangul Gumuz, Gambela, Harar, and Dire Dawa regions showed an increasing trend of first birth before age 20 in all the three surveys. In the same period the proportion of first birth before age 20 in Oromiya region decreased by 27.4%. Similarly, the proportion of first birth before age 20 differed by type of place of residence: urban and rural. Accordingly, from 2000 to 2011 the highest number of women who gave their first birth before age 20 resided in rural areas.

The proportion of women who gave their first birth before age 20 varied by occupational status. For instance, 37.3% of women were not working in 2000 at the time of survey. This proportion increased in 2005 year (71.3%) but decreased in 2011 (55.9%).

Additionally, we have pointed out the likelihood of having first birth before age 20 in Table 4.2 appendix A. The overall likelihood of having a first birth before age 20 increased from 36.84% in 2000 to 39.41% in 2005 and decreased slightly to 38.40 in 2011.

While there was a downward trend in the likelihood of first birth before age 20 among residents of the Tigray, Amhara and Addis Ababa, the likelihood increased for Affar and Oromiya from 2005 to 2011. The same likelihood decreased for other regions from 2005 to 2011.

Differentials were observed in the probability of having first birth by women's educational status. Among the different categories of women's education, having secondary and above education had the highest effect on the probability of having a teenage first birth compared with other educational attainments. More specifically, the likelihood of having first birth before age 20 was increasing for those women with no education (a trend that has continued over time) in the three surveys.

From 2000 to 2011 the likelihood of having first birth before age 20 for those living in urban areas decreased by 13.70 percent. On other hand, in 2005 and 2011, the likelihood of first birth before age 20 among women who lived in rural areas was at least three times of their counter parts that lived in urban areas.

Mass media exposure also appears to have effect on teenage birth. Women who were exposed to one of the three media (reading newspaper or magazine, listening radio, watching television) were less likely to have a first birth before age 20 than those of who were not exposed. But the chance of having first birth before age 20 for those women who had no exposure to any mass media was increasing over time (Table 4.2 appendix A).

Muslim women showed higher likelihood of having first birth before age 20 compared to women from other religions. More specifically, the probability of giving birth among teenage Muslims and Protestants increased slightly from 2000 to 2011, whereas the same likelihood for Coptic orthodox religious groups decreased by 25.5% during the same period.

Table 4.2 also showed that the likelihood of having first birth before age 20 for those women engaged in agricultural activity was higher than that of their counter parts. However, the likelihood decreased over time for those women involved in agricultural activity. The chance of teenage birth before age 20 for those women who were not working at time of interview increased in 2005 and slightly decreased in 2011.

Another description of the distribution event used a life-table (Table 4.3, Appendix A) that illustrated the key components of the population hazard functions amongst the sample of women in the three surveys. The first column gives the age of women at first birth. The next three columns tally the number of women who did not give first birth at the beginning of each full year, the number who gave first birth at the age and the censored numbers. According to 2000 EDHS, 1,435 women had their first birth before age 20. Similarly, for 2005 and 2011 EDHS the figures were 1,361 and 1,466 respectively. For the same periods 4,248(77.16%), 4,281(75.88%) and 5,136(77.80%) were censored (did not give birth at the time of the interview).

The fifth column of Table 4.3 presented another summary – the proportion of women who gave first birth by the end of each full year. We note that among the 6,283 women, 2.94 % had their first birth at age 15 in 2000 survey and increased to 4% in 2005 and dropped down to 3.68% in the 2011 survey. Of the 2,453 women who did not give first birth by age 18, 11.25% gave their first birth at age 19 in the 2011 survey.

Under the assumption of independent censoring we can use the sample hazard function to estimate the sample survival function at those ages when censoring precludes its direct computation. For example, an estimate of the survival probability at the end of age 18 is 0.8418*(1-0.0890). In other words, the sample survival probability in any year is simply one minus the hazard probability for that year multiplied by the sample survival probability from the previous year. Accordingly, the sixth column of Table 4.3 presented the proportion of women who did not give first birth at the end of each full age. Examining this sample survival function showed that 97% of the women did not give first birth at age 15 year in 2000 survey. For the same age, 96% of the women did not give first birth in 2005.This figure has slightly increased in 2011 to 96.32%.

Plotting the population hazard function illustrates the hazard experienced by women in each time period (Figure 1). According to the left panel of the Figure 1, there was no clear difference of first birth probability at age 17 in the 2000 and 2005 surveys. We can also see that there was no visible difference in the survival (probability of not giving first birth) at age 17 in the 2000 and 2011 surveys (the right panel). But the survival in the 2011 survey was higher than in the two previous surveys after age seventeen. We also note that the survival in the 2005 survey was lower than at the other two surveys for all the years 15-19.

### 4.2 Results of Discrete-Time Hazard Models

#### 4.2.1 Results of Baseline Model

In order to fit the model, we need to convert our data set, from what we refer to as a person-level data set, which contains one record for each person in the study, to a person-period data set, which contains one record for each time period that an individual is at risk of giving first birth before age 20. Table 4.4 illustrates the conversion from a person-oriented data set to a person-period data set using three individuals as a demonstration for the sample data from 2011 survey. A similar procedure was followed for 2000 and 2005 survey. The first two individuals have known their age at first birth — the first women gave first birth at age 18, and the second women gave first birth at age 17. The third woman (ID5560) had not yet given birth, so was censored at the end of age

19. The person-oriented data set describes the women's event histories using two variables : an event time (here DURATION, the period in which the individual experienced a first birth or was censored) and a censoring indicator (CENSOR = 0 for individuals who gave first birth and 1 for individuals who did not) and one time-invariant covariate, educational attainment of women (variable name: Education). The person-period data set includes a period variable, PERIOD, which specifies the time period t that the record describes. The particular time period described in the record is also identified through the set of time (age at first birth) indicator variables (A1 through A5).

Finally, the person-period data set includes an event indicator, $y_{it}$ which indicates whether a first birth occurred at period t (0 = no, 1 = yes). For each person, the event indicator must be 0 in every record except the last. Non-censored individuals (like individual 3686) experience the event in their last period, whereas censored individuals never experience a first birth, so $y_{it}$ remains 0 for all of their records (like ID5560).

Using Equation 11 (Chapter three, Section 3.3.2), the first model to be fitted was a simple discrete-time hazard model with no predictor and only a set age dummy variable (A1 through A5 described in Table 4.4). That is a baseline model. Using dummy variable A1 to A5 Equation 11 becomes:

$$\psi_{it} = \sum_{t=15}^{19} \alpha_t (AGE_{it}) = \alpha_1 A1 + \alpha_2 A2 + \alpha_3 A3 + \alpha_4 A4 + \alpha_5 A5$$ where

$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ stands for $\alpha_{15}, \alpha_{16}, \alpha_{17}, \alpha_{18}, \alpha_{19}$ respectively. Table 4.5 gives the parameter estimates for Model 1 in Equation 11.

The estimates found in Table 4.5 were the parameter estimates for the time-indicator variables (A1 though A5). These time indicator variables allowed for the estimation of the risk of event at each year (from 15 to 19 year). Accordingly, the estimated $\hat{\alpha}_1$ through $\hat{\alpha}_5$ describe the shape of the overall fitted logit-hazard profile. That is, if the risk of event occurrence were unrelated to time, the hazard function would be flat and the $\hat{\alpha}s$ are approximately equal. If event risks increased overtime, values of the $\hat{\alpha}s$ for latter periods would be greater than for earlier periods which is identical to estimated hazard in Table 4.5. For example, in 2000 EDHS (column 3), at age 15 we had $\hat{\alpha}_1 = -3.4952$ (s.e. =0.0746) and the estimate of $\alpha_1$ gave an estimate of hazard ($h_1$) to be $\hat{h}_1 = 0.0294$ (column 6). Women had a risk of 2.94 percent of giving first birth at age 15 in 2000. This figure increased to 4.01 percent in 2005 and then slightly decreased to 3.68 percent in 2011. For age 16, we had $\hat{\alpha}_2 = -2.8328$, $\hat{\alpha}_2 = -2.6872$, $\hat{\alpha}_2 = -2.8630$, in 2000, 2005 and 2011, respectively. Therefore women had 5.56, 6.37, and 5.4 percent risks of giving first birth at age sixteen in 2000, 2005 and 2011, respectively. Women had 9.6 percent risk of giving first birth at age 18 in 2000, 9.34 percent in 2005 and 8.89 percent in 2011. Similarly, the risk of having birth at age 19 was 12.08 % in 2000 and decreased to 11.25 in 2011. This indicated that the probability having birth at age 18 and 19 is decreasing over time.

### 4.2.2 Results of Discrete-time hazard model with covariate

The second model in this study was a discrete-time hazard model with demographic covariates (model 2 in Equation 12, Chapter three (Section 3.3.2)). But before fitting this model a univaraite discrete-time hazard model fit of each predictor variable against the response was performed to select significant candidate predictor variables that would qualify for the multivariable discrete-time hazard model. The result of univariate was shown in Table 4.6. According to Table 4.6 all covariates in this study were significant at 5%. That is, in a univariate discrete-time hazard model all variables in our study were significant at 5% and we included all variables that were significant in a univariate analysis for multivariable discrete-time hazard model (Table 4.7 appendix A). The results of multivariable discrete-time hazard model showed that four variables (region, education, place of residence, media) and time indicator (A1 through A5) variables were significant covariates at 5% in all three surveys (Table 4.7). A discrete-time hazard model containing four predictor (region, education, place of residence, media) were fitted for the three surveys and a model containing the predictors (region, religion, education, place of residence, occupation, and media) was fitted for the three surveys. The full model was written as:

$$\psi_{it} = \alpha_1 A1 + \alpha_2 A2 + \alpha_3 A3 + \alpha_4 A4 + \alpha_5 A5 + \beta_1 REG_i + \beta_2 EDU_i + \beta_3 PLR_{3_i} + \beta_4 MED_i$$
$$+ \beta_5 REL_i + \beta_6 OCC_i$$

And the reduced model was:

$$\psi_{it} = \alpha_1 A1 + \alpha_2 A2 + \alpha_3 A3 + \alpha_4 A4 + \alpha_5 A5 + \beta_1 REG_i + \beta_2 EDU_i + \beta_3 PLR_i + \beta_4 MED_i$$

Table 4.8 compares the full model and reduced model. In the reduced model region, education, place of residence, and exposure to media and the full model containing religion and occupation of women in addition to the other four were included. As shown in Table 4.7 the variables religion and occupation were not significant at 5% to indicate the outcome variable (first birth before age 20). According to Table 4.8 the difference of -2logL is insignificant in the three surveys and the four variables (region, education, place of residence, and exposure to media) were used for the rest of our analysis.

Accordingly, model 2 of Equation 12 becomes:

$$\psi_{it} = \sum_{t=15}^{19} \alpha_t (AGE_{it}) + \beta X_i$$
$$= \alpha_1 A1 + \alpha_2 A2 + \alpha_3 A3 + \alpha_4 A4 + \alpha_5 A5 + \beta_1 REG + \beta_2 EDU + \beta_3 PLR + \beta_4 MED \qquad (19)$$

where REG = region in which a woman lived , Edu =educational level of woman, PLR= place of residence of woman and MED=indicates a woman's expose to media. In the subsequent discussion we look for alternative parameterization of $\alpha_1 A1 + \alpha_2 A2 + \alpha_3 A3 + \alpha_4 A4 + \alpha_5 A5$ in Equation (19).

In Table 4.5 we provide estimates of parameter $\alpha_1\ to\ \alpha_5$ which represented the population hazard probability in each time period under consideration. There we showed that the hazard of first birth before age 20 was an increasing function in time. Parameterization of the hazard profile using time indicators (A1 through A5, in this case) lacks parsimony and representation of the main effect of time requires the inclusion of many parameters in the discrete-time hazard model.

Therefore, we need to adopt a particular algebraic form for the shape of the logit-hazard profile ($\alpha_1 A1 + \alpha_2 A2 + \alpha_3 A3 + \alpha_4 A4 + \alpha_5 A5$). A linear, quadratic, and cubic function of time (age at first birth or the indicator PERIOD in Table 4.4) were fitted. The plots of all functions were shown in Figure 2 for the 2011 survey data. Table 4.9 compares the deviance of all function. The general function is the full model $\alpha_1 A1 + \alpha_2 A2 + \alpha_3 A3 + \alpha_4 A4 + \alpha_5 A5$.

Because each lower order model is nested within each higher order model, values of the deviance statistic can be directly compared to help make analytic decisions. The difference in log likelihood value (-2 Log Likelihood) between the cubic and quadratic was 0.442=10615.21-10614-77 with p value=0.506

This indicated that quadratic function of age was a better fit than the cubic. Comparing linear and quadratic, the difference (-2Log likelihood) was 4.550 with p-value =0.0329. Therefore, we can fit a baseline hazard function with quadratic function of age at first birth. The same procedure was followed for 2000 and 2005 survey and a baseline hazard function

($\alpha_1 A1 + \alpha_2 A2 + \alpha_3 A3 + \alpha_4 A4 + \alpha_5 A5$) was fitted by quadratic function of age at first birth. Therefore equation (19) becomes:

$$\psi_{it} = Age + Age\_squared + \beta_1 REG_i + \beta_2 EDU_i + \beta_3 PLR_i + \beta_4 MED_i \qquad (20)$$

where Age = age at first birth of woman and Age_squared =square of age at first birth

Table 4.10 provides model comparison for Equations (19) and (20). Accordingly, -2logL was found to be insignificant in all the three survey, so that our final model became Equation (20).

Table 4.11 presents estimated coefficient, standard error, and hazard of timing of first birth before age 20 among women age 15-24 years at time of interview for model (20).

According to Table 4.11, the effect of region on the hazard of timing of teenage birth was not consistent across different regions and over time. For instance, in Tigray region the effect was significant for 2000 and 2005 data. That is, controlling the effect of other variables in the model, women who lived in Tigray in 2000 had a 25.6 %(HR=1.256) higher risk of having a teenage first birth compared to a women living in Oromiya region. The figure was 40.8% in 2005. But in 2011 there was no significant difference in risk between women who lived in Tigray and Oromiya. In 2000 and 2005 the hazard of first birth before age 20 of women who lived in Affar was not significantly different from that in Oromiya but the effect was significant in 2011 data. That is, after controlling the effect of other variables in the model, women living in Affar had 21.8% lower hazard of having a teenage birth compared to women living in Oromiya. The hazard of having first birth before age 20 for women living in Somali and Dire Dawa was not significantly different from those women living in Oromiya in 2005 and 2011. In 2005 and 2011 women who lived in SNNP had 83.8% higher risk of having a teenage first birth and the percentage decreased to 56.3% in 2011 compared to women who lived in Oromiya. The hazard of having first birth before age 20 was less in Addis Ababa compared to Oromiya. In 2000, 2005 and 2011 women who lived in Addis Ababa were 66.9%, 38.5% and 68.8% less likely to have a teenage birth compared women who lived in Oromiy respectively. On the other hand, the hazard of timing of first birth before age 20 was significantly lower in the three survey for Amhara region compared to Oromiya when the effect of other factors were controlled in the model.

Women with secondary and above level of education had a lower chance of having a birth at teenage compared to women who had no education and the decrease was more pronounced in 2011. In 2000 a woman with secondary and above level of education was 58.6% less likely to have first birth before age 20 compared to women with no education and this figure was increased to 72.8% and 85.4% in 2005 and 2011, respectively. In 2000, a woman with primary education was 25.5% less likely to have a teenage birth compared to women with no education and the figure was raised to 40.7% and 52.6% in 2005 and 2011 respectively. This indicated that the level of risk of having first birth before age 20 decreased from 2000 to 2011 for those women with primary, secondary and above schooling compared to those with no education.

Place of residence of women had a significant relationship with the hazard of timing of first teenage birth. Women who lived in rural areas were 40.9%, 53% and 39.8% more likely to have first birth before age 20 in 2000, 2005 and in 2011, respectively, controlling for others variables.

Women who had exposure to media were less likely to have early birth compared to those who had no exposure to mass media. Controlling for other variables in the model, women who had exposure to media in 2000, 2005 and 2011 were 28.6%, 17%, 15.8%, respectively, less likely to have a teenage birth compared to women with no expose to mass media.

### 4.3 Assumptions of discrete-time hazard models

Like all statistical models, the basic discrete-time hazard model invokes assumptions about the population that may, or may not, hold in practice. Therefore, before drawing conclusions on the final model, we verify three important assumptions of the discrete-time hazard model: the proportionality assumption, the linearity assumption and the no unobserved heterogeneity assumption (Singer and Willet, 1993).

1. The linearity assumptions: "all predictor operate only as main effects"
2. The proportionality assumption: "the effect of each predictor is constant"
3. No unobserved heterogeneity assumptions: "population hazard function depends only on predictor values".

The linear additivity assumption, similar to the linearity assumption in linear regression, implies that a predictor's effect does not depend on the value of another predictor in the model. This assumption can easily be tested by looking for interactions between predictors, and comparing the resulting model fits using deviance. Since we have no continuous predictors in this study, we don't have to bother about this assumption.

In testing the proportionality assumption, one tests whether each predictor in the model has the same effect in every time period under study. That is, an effect of covariates on the log-odds of having first birth before age 20 is the same at all time points. This assumption can be assessed by implementing interaction terms between the predictor and the quadratic function of time shown in Table 4.12.

Accordingly, in our study, every interaction between the time variable and predictors was considered, but the interaction was insignificant (Table 4.12) indicating poor model fits. Therefore, there is no reason to believe that the proportionality assumption has been violated.

Some individuals will be more at risk of experiencing an event than others, and it is unlikely that the reasons for this variability in the hazard will be fully captured by covariates. That is, the presence of unobserved (or unobservable) individual-specific risk factors leads to unobserved heterogeneity in the hazard. For checking this assumption, Singer and Willett (2003) recommend focusing on the hazard function, since unobserved heterogeneity has a consistent effect on the time variables and will lead to decreasing hazard functions. More specifically, whether the hazard of individual in a population is constant (or even increasing), the observed form of the hazard function at aggregate population level will tend to be decreasing and it seems that there are individual-specific unobserved factors. In this study, the hazard of individuals (probability of having first birth) is increasing over time and simultaneously the hazard function at aggregate population is increasing from age 15 to 19 (Table 4.3). Accordingly, we can say that unobserved heterogeneity seems unproblematic.

### 4.4 Model Diagnostics

We have checked the assumption of discrete-time survival analysis. We need to assess for influential observations. Accordingly we found out Cook's distance and DFBETA in Table 4.13. Cook's distance is a measure of the influences of cases. It is a measure of how much the residual of all cases would change if a particular case is excluded from the computation of the regression coefficients. The analog of Cook's influence statistic of a case is less than one indicates absence of potential outliers (Hosmer and Lemeshow, 2000) shown in Table 4.13 for 2011 data.

The maximum value of Cook's and DFBETA for each indicators of teenage first birth before age 20 less than one. This result indicated that among the indicators of teenage first birth before age 20, there are no potential influential observations (see Table 4.13). The same procedure was followed for 2000 and 2005 survey and in these two survey data there were no potential influential observations (Table 4.14 and 4.15, Appendix A).

# CHAPTER FIVE

## Discussion, Conclusion and Recommendations

### 5.1 Discussion

This study found evidence that some socio-economic variables had significant influence on first birth of women before age 20. Place of residence, educational status of women, region and media were found to be important determinants of first birth before age 20 among women aged (15-24 years).

Education was the covariate that showed a strong association with delayed childbearing among women of age 15-24 in Ethiopia in all three surveys. While primary education had a slight effect to delay first birth compared to no education, a strong effect was observed when women had secondary and higher education. This finding reinforces our understanding that women should be encouraged to complete secondary and higher education.

The result of this study agreed with the findings of studies done in Brazil (Gupta and Leite, 1999) and in other eight sub-Saharan African countries (Gupta and Mahy, 2003) which showed that education was found to be strongly associated with delayed childbearing among adolescents. More specifically, Gupta and Mahy (2003) found that education (grade 8 and above) consistently and significantly helped to reduce the risk of having first birth before age 20.

A study undertaken in Bangladesh about the relationship between adolescent motherhood and education by Kamal (2012) also agreed with our results. According to Kamal (2012), women's education, especially having secondary and higher level of education was an important determinant in delaying birth among women in Bangladesh. Another study in Bangladesh by Nahar and Min (2008) showed that attaining higher level of education has a strong delaying effect on first birth before age 20.

A study in Sweden by Olausson et al (2001) also revealed that teenage birth was positively associated with low educational attainment, which agreed with the results of this study. A study in Latin American countries by Elisa et al (2001) also agreed with this result. According to Elisa et al (2001) education showed a strong negative effect on teenage birth, especially up to 11-13 years of education. That is, the risk of having first birth before age 20 among women with 11-13 years of education was lower than the risk observed among those women with 0-3 years of education in the six Latin American countries under study (Bolivia, Brazil, Colombia, Guatemala, Dominican Republic and Peru).

Exposure to media was another important factor in explaining first birth before age 20. In the 2000 survey the risk of having first birth before age 20 for those women exposed to media was 28.6% less likely comparing to those who were not exposed to any media. The result decreased to 17.1% in 2005 and further decreased to 15.8% in 2011 survey (Table. 4.11). This could be because the huge increase in mass media exposure is a recent phenomenon and it takes time for media exposure to bring about increase in knowledge about teenage births. A study by Gupta and Mahy (2003) conducted in eight sub-Sahara African countries agreed with the result of this study. They found that regular radio listening habits was inversely associated with the probability of an adolescence first birth in Cotedevoire and Zimbabwe.

In this study place of residence was another socio-economic characteristic that was found to be associated with adolescent motherhood showing that urban women had lower proportion of teenage birth than rural women of the age group 15-19. Singh (1998), Gupta and Mahy (2003), Chandrasekhars (2010), and Kamal (2012) came up with similar conclusions. A reason for that could be that women in urban areas had better access to health services than those living in rural places (World Bank, 2004). On the other hand, results from studies done in Brazil and Colombia indicated that place of residence had no significant effect or loses its effect when the disparity in socio-economic levels was controlled (Gupta and Leite, 1999, Cesare and Rodriguez, 2006). Their result does not agree with result of this study which showed that when socio-economic variables are controlled early birth was low among teenagers living in urban areas in all the three surveys. A study in Nicaragua by Katherine et al (2009) in Nicaragua agreed with the result of this study which found that rural respondents were at higher risk of having an earlier first birth than were urban respondents.

According to the descriptive statistics the level of having first birth before age 20 was lower in Tigray, Affar, Amhara, Somali, Benishangul-Gumuz, SNNP, Gambela, Harari, Addis Ababa, and Dire Dawa compared to Oromiya in the three surveys. The results from discrete-time hazard model, however, showed that when socio-economic variables are controlled, the risk of teenage birth was higher in Tigray and Amhara than in Oromiya in the 2000 EDH survey. In 2005 survey data the risk of teenage birth was higher in Tigray, Amhara, Somali, Benishangul-Gumuz, SNNP, and Harari than in Oromiya controlling other variables in the model. The risk of teenage birth was higher in Benishangul-Gumuz and SNNP than in Oromiya in the 2011 survey.

### 5.2 Conclusion

In this study, we were not only interested in 'whether' a woman had first birth before age 20, but also in 'when' they had first birth (age at first birth). The study showed that 22.8%, 24%, and 22.2% of women aged 15-24 at time of each survey had first birth before age 20 in 2000, 2005, and 2011, respectively. The result of baseline model revealed that having first birth at age 15,16,17,18, and 19 was increasing in all surveys (Table 4.3). The study based on the discrete-time hazard model, revealed that residing in urban areas and having secondary and above level of education were the strongest effects in delaying first birth before age 20 in the three surveys. Exposure to media did not show an increment effect towards reducing teenage birth over time.

### 5.3 Recommendations

- The risk of having first birth before age 20 can be forestalled if we educate women. Therefore, working in the specialty of female education, particularly secondary and above level of education could minimize the risk of teenage birth.
- The hazard of having first birth before age 20 was higher among women residing in rural areas. As a result, concerted efforts are needed to empower adolescents to encourage the utilization of family planning targeting the rural teenagers.

## References


1. Alan Guttmacher Institute (AGI, 1998). Into a New World: Young Women's Sexual and Reproductive Lives, New York.
2. Allison. P.D. (1982). Discrete-Time Methods for the Analysis of Event Histories. *American Sociological Association*, **13**, 61-98.
3. Bicego, G. and J.T. Boerma (1993). Maternal and Child Survival. A Comperative Study Data from 17 Countries. *Social Science and Medicine* **36**(9), 1207-1227.
4. Bongaarts, J. (1978). A Framework for Analyzing the Proximate Determinants of Fertility. *Population and Development Review*, **4**(1), 105-132.
5. Bongaarts, J. (1982). The Fertility-Inhibiting Effect of the Intermediate Fertility Variables. *Studies in Family Planning*. **13**(6/7), 179-188.
6. Brown, C.C. (1975). On the Use of Indicator Variables for Studying the Time Dependence of Parameters in a Response-Time Model. *Biometrics*, **31**, 943-947.
7. Buvinic, Mayra (1998). The costs of adolescent childbearing: Evidence from Chile, Barbados, Guatemala, and Mexico. *Studies in Family Planning* **29**(2), 201-209
8. Byar, D. P., and Mantel, N. (1975). Some interrelationships among the regression coefficient estimates arising in a class of models appropriate to response-time data. *Biometrics* **31,** 943-947.
9. Central Statistical Agency of Ethiopia and ORC Macro. 2001. Ethiopia Demographic and Health Survey 2000. Addis Ababa, Ethiopia and Calverton, Maryland, USA:Central Statistical Agency and ORC Macro.

10. Central Statistical Agency of Ethiopia and ORC Macro. 2006. Ethiopia Demographic and Health Survey 2005. Addis Ababa, Ethiopia and Calverton, Maryland, USA:Central Statistical Agency and ORC Macro.
11. Central Statistical Agency of Ethiopia and ORC Macro. 2012. Ethiopia Demographic and Health Survey 2011. Addis Ababa, Ethiopia and Calverton, Maryland, USA:Central Statistical Agency and ORC Macro.
12. Cesare, M., and Rodríguez, J.V (2006). Micro analysis of adolescent fertility determinants: the case of Brazil and Colombia Papeles de POBLACIÓN No. 48 CIEAP/UAEM.
13. Chandrasekhar,S. (2010). Factors affecting age at marriage and age at first birth in india. *Journal of Quantitative Economics*, **8**(2), 81-97
14. Clark, S. (2004). Early marriage and HIV risks in sub-Saharan Africa. Stud Fam Plann, **35**,149-60.
15. Clark, S., Bruce. J, and Dude. A. (2006). Protecting young women from HIV/AIDS: the case against child and adolescent marriage. Int Fam Plan Perspect **32**, 79-88.
16. Cox, DR(1972). Regression Models and Life Tables. Journal of the Royal Statistical Society, Series B, **34**, 187-202
17. Davis, K. and J. Blake (1956). Social Structure and Fertility.An Analytical Framework. *Economic Development and Cultural Change*, **4**(2):211-235
*18.* Davis.WL, Olson.KW and Warner.L. (1993). An Economic Analysis of Teenage Fertility. Some Evidence from Oklahoma. *American Journal of Economics and Sociology,*52(1): 86-98
19. Dreze. J and M. Murthi. (2001). Fertility, education, and development: Evidence from India. *Population and Development Review*. **27**, 33-63.
20. Elisa.CF, Nunez.J. (2001).Teenage childbearing in Latin American Countries. Latin American Research Network, No-434
21. Gavin, L., C. Galavotti, H. Dube, A. D. McNaghten, M. Murwirwa, R. Khan, and M. St Louis. 2006. Factors associated with HIV infection in adolescent females in Zimbabwe. *J Adolesc Health* **39**(4)

22. Gill, K., pande.R, and Malhotra.A. ( 2007). Women deliver for development. *The Lancet* **370**, 1347-1357.
23. Gupta. N and Leite. IC.(1999) Adolescent fertility behavior. Trends and determinants in Northeastern Brazil. *International Family Planning Perspectives*, **25**(3),125-30.
24. Gupta. N and Mahy. M. (2003). Adolescent childbearing in sub-Saharan Africa. Can increased schooling alone raise ages at first birth? *Demographic Research,* 8(**4**), 93-106.
25. Hosmer, D. and Lemeshow, S. (2000). Applied Logistic Regression (2ndEdition). New York: John Wiley & Sons.
26. Islam,M. M. (1999). Adolescent Childbearing in Bangladesh. *Asia-Pacific Population Journal.* **14**(3), 73-87.
27. Kamal,M. (2012). Adolescent motherhood in Bangladesh: Evidence from 2007 BDHS data. *Canadian Studies in Population*, **39,** 63-82.
28. Katherine C. L, Prata.N and Stewart.C (2009). Adolescent Childbearing in Nicaragua: A Quantitative Assessment of Associated Factors. *International perspectives on Sexual and Reproductive Health,* 35(2), 91-96
29. Laird. N and Oliver.D.(1981). Covariance analysis of Censored Survival data Using log-linear analysis techniques. *Journal of the American Statistical Association*, **76**, 231-240.
30. Mantel. N and Hankey. B. (1978). A logistic regression analysis of response time data where the hazard function is time dependent." *Communications in Statistics-Theory and Methods,* **7** 333-347.
31. Mondal, N.I., Hossain, K. and Ali, K. (2009). Factors Influencing Infant and Child Mortality: A Case Study of Rajshahi District, Bangladesh, *J Hum Ecol*, **26**(1), 31-39.
32. Myers.M.H, Hankey. B.F, and Mantel.N. (1973). A logistic exponential model for use with response-time data involving regressor variables. *Biometrics* **29**, 257-269.
33. Nahar. Q and Min. H. (2008). Trends and determinants of adolescent child bearing in Bangladesh. Macro International Inc.

34. NIPORT (National Institute of Population Research and Training), Mitra and Associates, and ORC Macro. 2005. Bangladesh Demographic and Health Survey 2004. Dhaka, Bangladesh and Calverton, Maryland.
35. Nour, NM. (2006). Health consequences of child marriage in Africa. *Emerg Infect Dis* **12**(1), 1644-1649.
36. Olausson. P.O, Haglund .B, Weitoft .G.R, Cnattingius .S .(2001). Teenage childbearing and long-term socioeconomic consequences: a case study in Sweden. Department of Medical Epidemiology, Karolinska Institutet, Stockholm, Sweden.
37. Reardon.S.F, Brennan.R.T and Buka.S.L. (2002). Estimating Multi-Level Discrete-Time Hazard Models Using Cross-Sectional Data: Neighborhood Effects on the Onset of Adolescent Cigarette Use. Multivariate Behavioral Research. *Routledge*, **37**(3), 297-330
38. Schultz, T. P. (1981). Economics of population. Reading, Massachusetts. *Addison Wesley Publishing*, **20**(2), 150-190
39. Singer. J.D, Willett.J.B. (1993). It's about Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events. *Journal of Educational Statistics*, **18**(2), 155-195.
40. Singer. J.D and Willett.J.B. (2003). Applied Longitudinal Data Analysis. Modeling Change and Event Occurrence. *Oxford Univcrsity Prcss, Inc.*
41. Singh. S. (1998). Adolescent childbearing in developing countries. A global review. *Studies in Family Planning*. **29**(2), 117-136.
42. Singh.S, Darroch.E, Jennifer J. Frost and the Study Team (2001). Socioeconomic Disadvantage and Adolescent Women's Sexual and Reproductive Behavior: The Case of Five Developed Countries. *Family Planning Perspectives*, **33**(6), 251-258 and 289
43. Taffa.N and Obare.F (2004). Pregnancy and child health outcomes among adolescents in Ethiopia. *Ethiopian Journal of Health Development*, **18** (2), 90-95
44. Tewodros A., Jemal H., Dereje H.(2010). Determinants of adolescent fertility in Ethiopia. *Ethiop. J. Health Dev*. **24**(1), 30-38.

45. Tim.B. Heaton, Renata Forste, Samuel M. Otterstrom (2002). Family transitions in Latin America: first intercourse, first union and first birth. *International Journal of Population Geography*, **8**(1), 1-15.
46. Thompson, W. A., JR. 1977. On the treatment of grouped observations in life studies. *Biometrics*, **33**,463-470
47. United Nations Department of Economic and Social Affairs, Population Division, '2011 Update for the MDG Database: Adolescent birth rate', POP/DB/Fert/A/MDG 2011, reanalyzed by UNICEF
48. United Nations Children's Fund (UNICEF) (2012). Progress for Children. A report card on adolescents. No 10, 23-30
49. World Bank (2004) Ethiopia: A country status Reporton Health and Poverty, The World Bank Africa Region Human Development & Ministry of Health Ethiopia.
50. World Health Organization (2008). Department of Making Pregnancy Safer, Adolescent Pregnancy.**1**(1)
51. World Health Organization (2011). WHO Guidelines on Preventing Early Pregnancy and Poor Reproductive Outcomes among adolescents in Developing Countries.

# Appendix

## Appendix A: Tables

**Table 4.1: Percentage distribution of women who gave their first birth before age 20 by measure of socio economic characteristics for the three EDH Surveys**

| Variables | EDHS 2000 | EDHS 2005 | EDHS 2011 |
|---|---|---|---|
| | % | % | % |
| First birth before age 20 | 22.8 | 24.0 | 22.2 |
| **Education** | | | |
| No education | 74.3 | 68.2 | 52.3 |
| primary | 17.6 | 23.1 | 42 |
| Secondary and above | 8.2 | 8.7 | 5.7 |
| **Place of residence** | | | |
| Urban | 18.7 | 17.9 | 18.8 |
| Rural | 81.3 | 82.1 | 81.2 |
| **Religion** | | | |
| Orthodox | 46.7 | 45.4 | 35.5 |
| Protestant | 12.1 | 14.9 | 17.2 |
| Muslim | 37.7 | 36.7 | 45.4 |
| Others | 3.6 | 3.0 | 1.9 |
| **Occupation** | | | |
| Not working | 37.3 | 71.3 | 55.9 |
| Agricultural worker | 41.5 | 15.2 | 19.5 |
| Non- agricultural worker | 21.3 | 13.5 | 24.5 |

| | | | |
|---|---|---|---|
| **Exposure to media** | | | |
| Not at all | 64.5 | 58.5 | 42.2 |
| Exposure | 35.5 | 41.5 | 57.8 |
| **Region** | | | |
| Tigray | 11.5 | 10.6 | 12.4 |
| Affar | 6.7 | 5.2 | 9.8 |
| Amhara | 17.2 | 17.4 | 12.7 |
| oromiya | 19.7 | 17.4 | 14.3 |
| somali | 4.8 | 4.8 | 7.0 |
| Benishangul-gumuz | 8.6 | 9.0 | 10.2 |
| SNNP | 9.3 | 11.9 | 9.8 |
| Gambela | 6.8 | 7 | 9.1 |
| Harari | 6.2 | 6.6 | 6.8 |
| Addis ababa | 5.3 | 5.5 | 2.7 |
| Dire Dawa | 3.8 | 4.7 | 5.2 |

Source: Own calculation based on the 2000, 2005, and 2011 EDHS reports.

**Table 4.2: probability of having a first birth before age 20 by selected characteristics among women aged 15-24 at time of interview for the three surveys.**

| | | 2000 | 2005 | 2011 |
|---|---|---|---|---|
| | | % | % | % |
| Overall | | 36.84 | 39.41 | 38.40 |
| Region | Tigray | 59.12 | 54.7 | 44.67 |
| | Affar | 57.57 | 46.15 | 51.56 |
| | Amhara | 66.39 | 58.37 | 39.61 |
| | Oromiya | 43.90 | 42.65 | 44.01 |
| | Somali | 42.59 | 62.70 | 61.71 |
| | Benishangul-Gumuz | 57.12 | 63.46 | 56.35 |
| | Gambela | 27.66 | 33.04 | 31.08 |
| | SNNP | 53.74 | 63.35 | 55.74 |
| | Harari | 36.15 | 39.06 | 32.05 |
| | Addis Ababa | 11.34 | 11.31 | 6.77 |
| | Dire Dawa | 18.26 | 29.98 | 28.23 |
| Woman's Educational Attainment | No education | 51.13 | 62.53 | 69.52 |
| | Primary | 33.30 | 34.35 | 34.57 |
| | Secondary and above | 13.15 | 12.21 | 8.41 |
| Place of residence | Urban | 19.64 | 17.14 | 16.95 |
| | Rural | 49.33 | 56.09 | 51.67 |
| Exposure to Media | No | 52.24 | 63.28 | 65.11 |
| | Yes | 26.01 | 26.06 | 28.18 |
| Religion | orthodox | 37.25 | 34.61 | 27.75 |
| | Protestant | 31.17 | 35.14 | 37.01 |
| | Muslim | 43.04 | 49.94 | 49.22 |
| | Others | 38.54 | 46.89 | 37.68 |
| Occupation | not working | 34.79 | 43.42 | 40.82 |
| | Agricultural-worker | 61.43 | 57.93 | 48.63 |
| | Non-agricultural worker | 25.21 | 21.79 | 26.55 |

**Table 4.3: Life table describing the distribution of event occurrence over time (age)**

| | Number of | | | | | | | | | Proportion of | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Women who not gave first birth at the beginning of each age. (b) | | | Women who gave their first birth during the age (a) | | | Were censored at age | | | women at the beginning of the age who gave their first birth by the end of the age (a/b) | | | All women who did not gave their first birth at the end of each age | | |
| Age | 00 | 05 | 11 | 00 | 05 | 11 | 00 | 05 | 11 | 00 | 05 | 11 | 00 | 05 | 11 |
| 15 | 6283 | 5632 | 6602 | 185 | 226 | 243 | 826 | 683 | 879 | 0.0294 | 0.0401 | 0.0368 | 0.9706 | 0.9600 | 0.9632 |
| 16 | 5272 | 4723 | 5480 | 293 | 301 | 296 | 718 | 615 | 733 | 0.0556 | 0.0637 | 0.0540 | 0.9166 | 0.8987 | 0.9112 |
| 17 | 4261 | 3807 | 4451 | 350 | 316 | 339 | 567 | 516 | 603 | 0.0821 | 0.0830 | 0.0762 | 0.8413 | 0.8241 | 0.8418 |
| 18 | 3344 | 2975 | 3509 | 321 | 278 | 312 | 656 | 675 | 744 | 0.0960 | 0.0934 | 0.0890 | 0.7605 | 0.7471 | 0.7670 |
| 19 | 2367 | 2022 | 2453 | 286 | 230 | 276 | 2081 | 1792 | 2177 | 0.1208 | 0.1137 | 0.1125 | 0.6687 | 0.6621 | 0.6806 |

00, 05 and 11 stands for year 2000, 2005 and 2011 respectively

**Table 4.4 Conversion of a person-level data set into a person-period data set.**

"Person-level data set"

| ID | DURATION | CENSOR | Education |
|---|---|---|---|
| 3686 | 18 | 0 | 0 |
| 5440 | 17 | 0 | 2 |
| 5560 | 19 | 1 | 1 |

"Person-period data set"

| ID | PERIOD | A1 | A2 | A3 | A4 | A5 | Education | $y_{it}$ |
|---|---|---|---|---|---|---|---|---|
| 3686 | 15 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3686 | 16 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 3686 | 17 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 3686 | 18 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 5440 | 15 | 1 | 0 | 0 | 0 | 0 | 2 | 0 |
| 5440 | 16 | 0 | 1 | 0 | 0 | 0 | 2 | 0 |
| 5440 | 17 | 0 | 0 | 1 | 0 | 0 | 2 | 1 |
| 5560 | 15 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 5560 | 16 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 5560 | 17 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 5560 | 18 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 5560 | 19 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |

**Table 4.5 Parameter Estimates and Fitted Hazard Probabilities from baseline discrete-time hazard model fitted to data 2000, 2005, and 2011 EDHS**

| | | Parameter Estimate ($\hat{\alpha}$) and Standard errors (s.e) | | | fitted hazard | | |
|---|---|---|---|---|---|---|---|
| Period | predictor | 2000 | 2005 | 2011 | 2000 | 2005 | 2011 |
| 15 | A1 | -3.4952*** | -3.1747*** | -3.2645*** | 0.0294 | 0.0401 | 0.0368 |
| | | (0.0746) | (0.0679) | (0.0654) | | | |
| 16 | A2 | -2.8328*** | -2.6872*** | -2.8630*** | 0.0556 | 0.0637 | 0.0540 |
| | | (0.0601) | (0.0596) | (0.0598) | | | |
| 17 | A3 | -2.4136*** | -2.4022*** | -2.4957*** | 0.0821 | 0.0830 | 0.0762 |
| | | (0.0587) | (0.0587) | (0.0565) | | | |
| 18 | A4 | -2.2426*** | -2.2723*** | -2.3270*** | 0.0960 | 0.0934 | 0.0889 |
| | | 0.0889 | (0.0630) | (0.0593) | | | |
| 19 | A5 | -1.9846*** | -2.0530*** | -2.0653*** | 0.1208 | 0.1137 | 0.1125 |
| | | (0.0681) | (0.0700) | (0.0639) | | | |

* * * p-value <0.001

**Table 4.6 Result of univariate discrete-time hazard model**

| Dummy variables (A to A5) + covariates | DF | 2000 | | 2005 | | 2011 | |
|---|---|---|---|---|---|---|---|
| | | Chi-square | P(>Chi-square) | Chi-square | P(>Chi-square) | Chi-square | P(>Chi-square) |
| Region | 10 | 483.2 | <0.0001 | 302.94 | <0.0001 | 351.76 | <0.0001 |
| Education | 2 | 316.29 | <0.0001 | 496.24 | <0.0001 | 66.52 | <0.0001 |
| Place of residence | 1 | 185.84 | <0.0001 | 288.97 | <0.0001 | 368.86 | <0.0001 |
| Media | 1 | 258.7 | <0.0001 | 284.61 | <0.0001 | 223.46 | <0.0001 |
| Religion | 3 | 21.63 | 0.001905 | 46.407 | <0.001 | 89.973 | <0.0001 |
| Occupation | 2 | 64.546 | <0.0001 | 12.283 | 0.002152 | 64.201 | <0.0001 |

**Table 4.7.Results of multivariate discrete-time hazard model**

| Covariates | DF | 2000 Wald Chisquare | 2000 P(>Chi-sqaure) | 2005 Wald Chisquare | 2005 P(>Chi-sqaure) | 2011 Wald Chisquare | 2011 P(>Chi-sqaure) |
|---|---|---|---|---|---|---|---|
| A1 | 1 | 333.6637 | <0.0001 | 251.582 | <0.0001 | 304.5293 | <0.0001 |
| A2 | 1 | 205.0668 | <0.0001 | 163.6932 | <0.0001 | 215.8535 | <0.0001 |
| A3 | 1 | 126.9186 | <0.0001 | 112.7526 | <0.0001 | 141.0144 | <0.0001 |
| A4 | 1 | 91.4218 | <0.0001 | 86.1860 | <0.0001 | 102.4603 | <0.0001 |
| A5 | 1 | 47.3914 | <0.0001 | 53.0427 | <0.0001 | 57.0220 | <0.0001 |
| Religion | 3 | 3.9388 | 0.2682 | 3.1308 | 0.3719 | 6.1683 | 0.1037 |
| Region | 10 | 174.7428 | <0.0001 | 110.9250 | <0.0001 | 128.8013 | <0.0001 |
| Education | 2 | 51.2598 | <0.0001 | 158.2179 | <0.0001 | 265.2444 | <0.0001 |
| Place of residence | 1 | 12.0990 | 0.0005 | 15.6421 | <0.0001 | 12.3117 | 0.0005 |
| Occupation | 2 | 5.486 | 0.0644 | 4.2947 | 0.11691 | 1.0045 | 0.6051 |
| Media | 1 | 22.1049 | <0.0001 | 11.396 | 0.0007 | 6.5276 | 0.0106 |

**Table 4.8 model comparison**

| | -2logL 2000 | -2logL 2005 | -2logL 2011 |
|---|---|---|---|
| Full model | 1011.560 | 9028.931 | 9722.757 |
| Reduced model | 10120.230 | 9031.712 | 9725.593 |
| p-value | 0.15961 | 0.24895 | 0.24220 |

**Table 4.9 Comparison with for the baseline model**

| Representation for age at first birth | Deviance | AIC |
|---|---|---|
| Constant | 10840.62 | 10842.62 |
| Linear | 10619.76 | 10623.76 |
| Quadratic | 10615.21 | 10621.21 |
| Cubic | 10614.77 | 10622.77 |
| general | 10612.37 | 10622.37 |

**Table 4.10 Model comparison for Equation 19 and 20**

| | -2logL<br>2000 | -2logL<br>2005 | -2logL<br>2011 |
|---|---|---|---|
| Original model(19) | 10120.230 | 9031.712 | 9725.593 |
| Reparametrized model(20) | 10124.163 | 9033.858 | 9727.059 |
| p-value | 0.14016 | 0.34198 | 0.48047 |

**Table 4.11: Estimated coefficient, standard error, and hazard ratio of timing of first birth before 20 from discrete-time hazard model for women aged 15-24 at time of interview for the three surveys.**

| Variables | 2000 | | 2005 | | 2011 | |
|---|---|---|---|---|---|---|
| | Coeff.(SE) | HR | Coeff.(SE) | HR | Coeff.(SE) | HR |
| Intercept | -4.5122(0.1966) | - | -3.6123(0.1905) | - | -3.2360(0.1862) | - |
| **Region** | | | | | | |
| Tigray | 0.2279(0.1015) | 1.256* | 0.3421(0.1136) | 1.408** | 0.1722(0.1080) | 1.188 |
| Affar | 0.0820(0.1240) | 1.085 | -0.1682(0.1452) | 0.845 | -0.2458(0.1188) | 0.782* |
| Amhara | 0.5032(0.0863) | 1.654*** | 0.2030(0.0990) | 1.225* | -0.2165(0.1077) | 0.805* |
| Oromiya (ref.) | - | 1.00 | - | 1.00 | - | 1.00 |
| Somali | -0.3396(0.1393) | 0.712* | 0.0612(0.1519) | 1.063 | 0.0459(0.1336) | 1.047 |
| Benishangul -Gumuz | 0.1638(0.1121) | 1.178 | 0.3613(0.1209) | 1.435** | 0.2450(0.1160) | 1.278* |
| Gambel | -0.6721(0.1063) | 0.511*** | -0.2541(0.1053) | 0.776* | -0.4389(0.1143) | 0.645*** |
| SNNP | 0.1701(0.1241) | 1.185 | 0.6089(0.1345) | 1.838*** | 0.4466(0.1205) | 1.563*** |
| Harari | -0.1534(0.1290) | 0.858 | 0.4408(0.1373) | 1.554** | 0.1898(0.1331) | 1.209 |
| Addis Ababa | -1.1047(0.1477) | 0.331*** | -0.4864(0.1528) | 0.615** | -1.1653(0.1901) | 0.312*** |
| Dire Dawa | -0.9291(0.1560) | 0.395*** | 0.2223(0.1528) | 1.249 | -0.1990(0.1456) | 0.820 |
| **Woman's educational attainment** | | | | | | |
| No education (ref.) | - | 1.00 | - | 1.00 | - | 1.00 |
| Primary | -0.2941(0.0784) | 0.745*** | -0.5225(0.0747) | 0.593*** | -0.7456(0.0638) | 0.474*** |
| Secondary and above | -0.8820(0.1175) | 0.414*** | -1.3037(0.1156) | 0.272*** | -1.9242(0.1286) | 0.146*** |
| **Woman's place of Residence** | | | | | | |
| Urban (ref.) | - | 1.00 | - | 1.00 | - | 1.00 |
| Rural | 0.3429(0.0865) | 1.409*** | 0.4273(0.1020) | 1.533*** | 0.3351(0.0861) | 1.398*** |
| **Exposure to media** | | | | | | |
| Yes | -0.3365(0.0693) | 0.714*** | -0.1879(0.0704) | 0.829** | -0.1717(0.0626) | 0.842** |
| No (ref.) | - | 1.00 | - | 1.00 | - | 1.00 |

*p<=0.05;**p<=0.01;***p<=0.001; SE: standard error; HR: hazard ratio; ref.: reference

**Table 4.12 interaction of covariates with time**

| Covariate Interacted with | | | |
|---|---|---|---|
| Age and Age_square | DF. | Chi-square | p(>Chi-square) |
| Education | 2 | 4.0548 | 0.1317 |
| Place of residence | 2 | 0.9795 | 0.6128 |
| Region | 2 | 2.8983 | 0.2348 |
| Media | 2 | 5.5021 | 0.0639 |

**Table 4.13 Summary of Influential Variables Statistics (for 2011 data)**

| Influential variables | cases | Minimm | Maximum |
|---|---|---|---|
| Analog of Cook's influence statistics | 22495 | .00000 | .00015 |
| Deviance residuals | 22495 | -.34823 | 1.22692 |
| DFBETA for constant | 22495 | -.00954 | .01022 |
| DFBETA for Tigray | 22495 | -.00608 | .00948 |
| DFBETA for Affar | 22495 | -.00554 | .00642 |
| DFBETA for Amhara | 22495 | -.00578 | .00609 |
| DFBETA for Somali | 22495 | -.00590 | .00807 |
| DFBETA for Benishangul-Gumuz | 22495 | -.00541 | .00784 |
| DFBETA for Gambela | 22495 | -.00620 | .00908 |
| DFBETA for SNNP | 22495 | -.00801 | .01221 |
| DFBETA for Harari | 22495 | -.01027 | .00559 |
| DFBETA for Addis Ababa | 22495 | -.00764 | .01191 |
| DFBETA for DireDawa | 22495 | -.00723 | .01276 |
| DFBETA for place of residence (urban) | 22495 | -.00734 | .00587 |
| DFBETA for education(primary) | 22495 | -.00318 | .00371 |
| DFBETA for education(secondary and above ) | 22495 | -.00504 | .01242 |

| | | | |
|---|---|---|---|
| DFBETA for exposure to media (yes) | 22495 | -.00339 | .00254 |

**Table 4.14 Summary of Influential Variables Statistics (for 2000 data)**

| | N | Minimum | Maximum |
|---|---|---|---|
| Analog of Cook's influence statistics | 21527 | .00000 | .000016 |
| Deviance value | 21527 | -.73655 | 1.20240 |
| DFBETA for constant | 21527 | -.0078 | .0092 |
| DFBETA for Tigray | 21527 | -.00675 | .01030 |
| DFBETA for Affar | 21527 | -.00687 | .00820 |
| DFBETA for Amhara | 21527 | -.00661 | .00726 |
| DFBETA for Somali | 21527 | -.00689 | .00936 |
| DFBETA for Benishangul-Gumuz | 21527 | -.00687 | .00838 |
| DFBETA for Gambela | 21527 | -.00816 | .01055 |
| DFBETA for SNNP | 21527 | -.00978 | .01157 |
| DFBETA for Harari | 21527 | -.01309 | .01063 |
| DFBETA for Addis Ababa | 21527 | -.01008 | .00134 |
| DFBETA for DireDawa | 21527 | -.00717 | .01004 |
| DFBETA for Place of residence (Urban) | 21527 | -.00894 | .00634 |
| DFBETA for exposure to Media(yes) | 21527 | -.00458 | .00443 |
| DFBETA for Education(primary) | 21527 | -.00411 | .00588 |
| DFBETA for Education(secondary and above) | 21527 | -.00751 | .01089 |

**Table 4.15 Summary of Influential Variables Statistics (for 2005 data)**

| | N | Minimum | Maximum |
|---|---|---|---|
| Analog of Cook's influence statistics | 19159 | .00000 | .000023 |
| Deviance value | 19159 | -.86218 | 1.14968 |
| DFBETA for constant | 19159 | -.01342 | .00591 |
| DFBETA for Tigray | 19159 | -.00909 | .01027 |
| DFBETA for Affar | 19159 | -.00809 | .00545 |
| DFBETA for Amhara | 19159 | -.00779 | .00546 |
| DFBETA for Somali | 19159 | -.00801 | .00999 |
| DFBETA for Benishangul-Gumuz | 19159 | -.00778 | .00769 |
| DFBETA for Gambela | 19159 | -.00906 | .01059 |
| DFBETA for SNNP | 19159 | -.01085 | .01084 |
| DFBETA for Harari | 19159 | -.01402 | .01080 |
| DFBETA for Addis Ababa | 19159 | -.01182 | .00937 |
| DFBETA for Dire Dawa | 19159 | -.00938 | .01081 |
| DFBETA for Place of residence(Urban) | 19159 | -.01083 | .00208 |
| DFBETA for Education(primary) | 19159 | -.00332 | .00487 |
| DFBETA for Education(secondary and above) | 19159 | -.00606 | .00265 |
| DFBETA for exposure to media(yes) | 19159 | -.00481 | .00420 |

## Appendix B: Figures

**Figure 1: Sample survivor (did not give first birth before age 20) and hazard (gave first birth before age 20) probability for all women aged 15-24 at the time interview in each survey.**

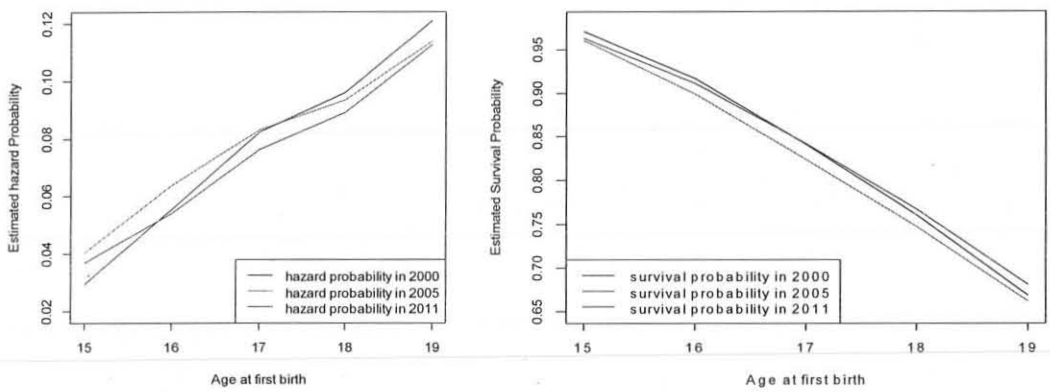


**Figure 2: Fitted functions for the baseline model for the 2011 EDHS data**