JIMMA UNIVERSITY
COLLEGE OF NATURAL SCIENCE
DEPARTMENT OF STATISTICS

Survival Analysis of Age at First Marriage and Age at First Sexual Intercourse among Women in Rural Ethiopia

By:
Yihenew Mitiku

June, 2016
JIMMA, ETHIOPIA

Survival Analysis of Age at First Marriage and Age at First Sexual Intercourse among Women in Rural Ethiopia

By:
Yihenew Mitiku

Advisor: Belay Berlie (PhD Scholar)
Co-advisor: Demeke Keflie (Msc)

June, 2016
JIMMA, ETHIOPIA

STATEMENT OF AUTHOR

I declare that this thesis is a result of my genuine work and all sources of materials used have been duly acknowledged. I have submitted this thesis to Jimma University in the partial fulfillment for the Degree of Master of Science in Biostatistics. The thesis can be deposited in the university library to be made available to borrowers for reference. I solemnly declare that I have not so far submitted this thesis to any other institution anywhere for that award of any academic degree, diploma or certificate. Brief quotations from this thesis are allowed without requiring special permission provided that an accurate acknowledgement of the source is made. Requisites for extended quotations for the reproduction of the thesis in whole or in part may be granted by the head of the department of statistics when in his or her judgment the proposed use of the material is for a scholarly interest. In all other instances, however, permission must be obtained from the author.

Yihenew Mitiku
Date: _____
Signature: _____
Jimma University, Jimma, Ethiopia

DEPARTMENT OF STATISTICS, SCHOOL OF GRADUATE STUDIES

JIMMA UNIVERSITY

As thesis research advisors, we herby certify that we have read the thesis prepared by Yihenew Mitiku, under our guidance, which is entitled "**Survival Analysis of Age at First Marriage and Age First Sexual Intercourse among Women in Rural Ethiopia**", in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including tables and figures are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

BELAY BERLIE(PhD Scholar)	_____	_____
Main Advisor	Signature	Date
DEMEKE KEFELIE(MSc)	_____	_____
Co- Advisor	Signature	Date

As the members of the board of examiners of MSc thesis open defense examination, we certify that we have read and evaluated the thesis and examined the candidate. Hence, we recommend that the thesis be accepted as it fulfills the requirements for the degree of Master of Science in Biostatistics.

_____	_____	_____
Name of Chairman	Signature	Date
_____	_____	_____
Name of Advisor	Signature	Date
_____	_____	_____
Name of Co-Advisor	Signature	Date
_____	_____	_____
Name of Internal Examiner	Signature	Date
_____	_____	_____
Name of External Examiner	Signature	Date

ACKNOWLEDGEMENT

First and foremost, I am indebted to the Almighty God because of whose full mercy and grace I could complete my study.

I would like to give deep gratitude to my advisor, Belay Berlie (PhD Scholar) without him accomplishment of this work would not have been made possible. His continuous supportive comments and constant guidance throughout the course of my thesis work get me out of my frustration. It is a real pleasure working with him. I hope that our cooperation will keep on going in the future. I gratefully acknowledge my co-advisor Demeke keflie(MSc) for his continuous guidance and support though out the study .

I also like to thank my wife Bertucan Ademe for being with me and helping me by giving constructing ideas throughout my study time.

I would like to thank the host Jimma University and my sponsor Assosa University for providing me to attend my training and Ethiopian Central Statistics Agency for providing me with all the relevant secondary data used in this study.

Finally, I would like to thank my parents who have invested all their life to hold up me, and I would like to thank my classmates for their support and knowledge sharing in my two years of study.

DEDICATION

This thesis is dedicated to my family and to my beloved wife Bertucan Adem.

ABSTRACT

Introduction: Age at first marriage and first sex has important implications for gender relations and the organization of family life in societies. Age at marriage is of particular interest because it marks the transition to adulthood in many societies; the point at which certain options in education, employment, and participation in society are foreclosed; and the beginning of regular socially acceptable time for sexual activity and childbearing. This study aimed to investigate the potential risk factors affecting age at first marriage and age at first sexual intercourse using Cox proportional hazards and accelerated failure time models among women in Rural Ethiopia.

Methods: The data source for the analysis was the 2011 Ethiopian Demographic and Health Survey data. The study considered 10,417 women aged 15-49 years from nine regions and one city administration. The Cox proportional hazards and accelerated failure time models were employed with the help of R statistical package and STATA soft wares.

Results: The median survival time for both age at first marriage and age at first sexual intercourse were 17 years with 95% CI; (16.90, 17.10) and (16.90, 17.11) respectively. Log-logistic accelerated failure time model was better than weibull and log-normal models based on Akaike's information criterion and graphical evidence. The result showed that region, women's educational level, wealth index and religion were significantly affect timing of first marriage and first sexual intercourse. Women who had secondary and higher education prolonged time-to-first marriage by the factor of ϕ =1.42 and ϕ =1.46, respectively, and prolonged time-to-first sexual intercourse by the factor of ϕ=1.38 and ϕ=1.34.

Conclusion: Improving girls and young women access to education is important for rising the women's age at first marriage and first sexual intercourse, which is vital for empowering them and enhancing their participation in any sector.

Key Words: Survival Data Analysis, Time to First Marriage, Acceleration Factor, Time to First sexual intercourse

ACRONYMS

AFM	Age at First Marriage
AFSI	Age at First Sexual Intercourse
AFT	Accelerated Failure Time
AIC	Akaike Information Criterion
CI	Confidence Interval
CSA	Central Statistics Agency
DHS	Demographic and Health Survey
EAs	Enumeration areas
EDHS	Ethiopian Demographic and Health Survey
HIV/AIDS	Human Immune deficiency Virus/ Acquired Immune deficiency Syndrome
HR	Hazard Ratio
KM	Kaplan-Meier estimate
MOFED	Ministry of Finance and Economic Development
NCTPE	National Committee on Harmful Traditional Practices of Ethiopia
NDHS	Nigeria Demographic and Health Survey
NRC-IOM	The National Research Council and Institute of Medicine of the National Academies
OR	Odd Ratio
Pdf	probability density function
PH	Proportional Hazards
PO	Proportional odds
SNNPR	Southern Nations, Nationalities and Peoples representatives
STDs	Sexually Transmitted Diseases
UN	United Nation
UNFPA	United Nations Fund for Population Activities
UNICEF	United Nations Children's Fund

Table of Contents

Statement of Author i
Acknowledgement iii
Abstract v
Acronyms vi
List of Tables ix
List of Figures ix
1 Introduction 1
1.1 Analysis of Time to Event Data 1
1.2 Age at First Marriage and Age at First Sexual Intercourse 2
1.3 Statement of the Problem 6
1.4 Objective of the Study 8
1.4.1 General Objective 8
1.4.2 Specific Objectives 8
1.5 Significance of the Study 8
2 Literature Review 9
2.1 Early Marriage 9
2.1.1 Age at First Marriage 9
2.1.2 Risk Factors of Age at First Marriage 10
2.2 Early Sexual Intercourse 16
2.3 Review of Survival Models 16
3 Methodology 18
3.1 Data 18
3.1.1 Source of Data 18
3.1.2 Sampling Design 18
3.1.3 Variables in the Study 19
3.2 Method of Data Analysis 20
3.2.1 Survival Analysis 20
3.2.2 Non-parametric Survival Methods 21
3.2.3 Cox Proportional Hazards Regression Model 22
3.2.3.1 Model Diagnostics 23

3.2.4 Accelerated Failure Time Model 25
3.2.4.1 Parameterization 29
3.2.4.2 Parameter Estimation of AFT Model 30
3.2.4.3 Model Selection 30
3.2.4.4 Model Diagnostics: AFT Model 31
4 Results 33
4.1 Exploratory Data Analysis 33
4.2 Cox Proportional Hazards Model 39
4.2.1 Model Diagnostics 41
4.3 Accelerated Failure Time Model 44
4.3.1 Univariable Analysis 44
4.3.2 Multivariable AFT Model 45
4.3.3 Model Diagnostics: AFT Model 48
4.4 Comparison of Cox PH and Log-logistic AFT Model 51
4.5 Discussion 52
5 Conclusions and Recommendations 57
5.1 Conclusions 57
5.2 Recommendations 58
References 59
Appendix 66
Appendix A 66
Appendix B 69
Appendix C 72

LIST OF TABLES

Table 3. 1: Coding and explanation of explanatory variables 19
Table 4. 1: Descriptive Summary of Demographic and Socio-economic Variables.....33
Table 4. 2: Median Time of First Marriage and First Sexual Intercourse result of Non-parametric test 38
Table 4. 3: Univariable Analysis of AFM and AFSI using Cox PH Model 39
Table 4. 4: Multi-variable Analysis of Age at First Marriage and First Sexual Intercourse using Cox PH Model 40
Table 4. 5: Cox Proportional Odds Assumption test for the variables included in the multivariable model 42
Table 4. 6: Comparison of AFT Models Using AIC criteria for AFM and AFSI data..... 45
Table 4. 7: Log-logistic Multi-variable AFT Model for AFM and AFSI data 46
Table 4. 8: Comparison of Cox PH and log-logistic AFT Models 52

LIST OF FIGURES

Figure 4. 1: The K-M plots of Survival and hazard function of age at first marriage 35
Figure 4. 2: The K-M plots of Survival and hazard function of age at first sexual intercourse ... 36
Figure 4. 3: Test for Proportional Odds Assumption for the Cox Model 43
Figure 4. 4: Cox- Snell residuals plots for AFM and AFSI data respectively 44
Figure 4. 5: Graphs of Weibull, Log-logistic, and Log-normal baseline distributions for age at first marriage and first sexual intercourse data sets 49
Figure 4. 6: Cox- Snell residuals plots of log-logistic baseline distribution for age at first marriage and first sexual intercourse data respectively 50
Figure 4. 7: Quantile- Quantile plot to check the adequacy of the accelerated failure time model 51

1 Introduction

1.1 Analysis of Time to Event Data

Analysis of time to event data arises more often in various areas of study, for instance in engineering where reliability of equipment is judged by time to failure or in clinical trials where interest is in the time until disease healing, pain cessation, or time till death in terminal diseases such as cancer. The latter case scenario is well reflected in the common terminology: 'Survival Data Analysis' arising from mortality studies. Data arising from such study designs are characterized by censoring which renders the standard statistical methods such as linear models and generalized linear models inapplicable. Further still, the population considered is often changing as only individuals at risk at a given time point are considered, an aspect known in survival data analysis as conditioning (Duchateau and Janssen, 2008). Only right censoring is considered here as it is the most relevant and also the most common in clinical trial settings. Right censoring occurs when an event fails to be realized within the stipulated study timeframe (Wienke, 2011). The other types of censoring are left and interval censoring; and are extensively discussed in (Getachew et al, 2009).

The Kaplan Meier curve (Kaplan and Meir, 1958) is the most common non-parametric estimation approach applied in survival data analysis. In this approach, the data are summarized in a plot of survival functions from which summary statistics such as median time to event or event free rate at a time point can be estimate (Legrand, 2005). The Kaplan Meier curve assumes independence and homogeneity between observations. However, in most cases, observations in Survival data analysis exhibit a form of heterogeneity (Korosteleva, 2009). For instance in clinical trials, patients may differ in terms of age, gender and many other characteristics. It is often of more interest to study the relationship between survival times and such covariates often measured at baseline particularly if they can be controlled. Though stratified Kaplan Meier curves can be applied, it is not capable to simultaneously take many such variables, into account. Furthermore, it is difficult to quantify the effect of the stratifying variable on the survival times (Fox, 2002). The Proportional hazards model introduced by Cox (1972) is a regression model with event time as the response variable. Unlike the Kaplan Meier approach, the model can

simultaneously take several covariates, often referred to as risk factors, into account. The Cox model is perhaps the most widely used in analysis of time to event data today. The model assumes a common study population risk of death or event (hazard) function known as the baseline hazard. Individual subject information is embedded in the model through the covariates acting multiplicatively on the baseline hazard increasing or decreasing individual risk based on the prognostic information (Wienke, 2011). The Accelerated Failure Time (AFT) model is another alternative method for the analysis of survival data. Many of the standard parametric models such as Weibull, Exponential, Lognormal and log logistic are accelerated failure time models (Kleinbaum D, 2005). Although the Cox regression model is the most favorable employed technique in survival analysis, parametric models (Lawless JF, 1998) do have a number of benefits.

1.2 Age at First Marriage and Age at First Sexual Intercourse

Birth, marriage and death are key events in most people's lives. But only marriage is the matter of choice. The National Research Council and Institute of Medicine of the National Academies (NRC-IOM, 2005), states that the right to exercise the choice of marriage has been established in international human rights instruments. According to human rights advocates, marriage before 18 years contravenes the United Nations Convention on the rights of the child, which defines age18 as the end of childhood. Therefore, marriage before 18 years can be considered as child marriage (Population Council, 2002). Moreover, very early marriage is said to undermine other rights guaranteed by the convention, including the right to be protected from physical abuse and sexual exploitation and the right not to be separated from parents against one's will (Population Council, 2002). Yet many girls, and a smaller number of boys, enter marriage without any chance of exercising their right to choose. Some are forced into marriage at a very early age; others are simply too young to make an informed decision about their marriage partner or about the implications of marriage itself.

Age at first marriage and first sexual intercourse has important implications for gender relations and the organization of family life in societies (Mensh BS et al., 2005). Studies in developed countries have found that the widespread availability of contraception is associated with a gradual decoupling of sexual debut and entry into marriage among women (Waite LJ, 2006).

When contraceptives are widely available, sexual intercourse tends to occur early, whereas marriage is delayed, leaving prolonged periods for sexual experimentation and mate selection (Goldin and Katz, 2000).

Age at first marriage is the time or age at which both male and female enter into marriage and their marital union is formed and assessed in years (United Nations, 2000). Age at marriage is of particular interest because it marks the transition to adulthood in many societies; the point at which certain options in education, employment, and participation in society are foreclosed; and the beginning of regular socially acceptable time for sexual activity and childbearing. As such marriage is not only the most predominant context for childbearing but also one of the most important determinants of fertility (Lesthaeghe et al., 1989). However, reproduction is not the only function of marriage. Marriage marks the beginning of a new family unit with all the complicated statuses and the roles that the members of this unit are expected to play. For the society as a whole, marriage unites several individuals from different families and represents the creation of a production and consumption unit as well as one for the exchange of goods and services (Ikamari, 2005).

Age at first marriage is often used as a proxy for first exposure to intercourse and risk of pregnancy. But the two events may not occur at the same time because some women may engage in sexual activity before marriage. Early sexual debut increases the risk of unprotected sexual intercourse, multiple partnership, and sexually transmitted infections (STIs), including HIV/ AIDS (Blanc,1998). Indeed, recent studies have examined early sexual activity largely as a potential risk factor for adverse social and health outcomes (Welz, 2007). Beginning sexual activity exposes an individual to the risk of acquiring sexually transmitted infections, especially if the age at debut is too early for the individual to have acquired the necessary sex education to navigate relationships successfully (Hallett, 2007). It is also reported that people who begin to have sex earlier than others of their age are more likely to be infected with HIV.

In sub-Saharan African countries, there is a wide range in age at first sexual intercourse as well as age at first marriage. The variation in the timing of first sex has been attributed to diverse environmental and social factors such as access to family planning programs, culture, and the prevalence of HIV/AIDs (Wellings et al., 2006). In general, there is an overall trend of

increasing age at first marriage (Wellings et al., 2006). Most evidence indicates that delayed marriage is associated with changes in women's status, especially increased school enrollment and employment opportunities (Kaufmann and Meekers, 1998).

In many countries, delayed marriage has been associated with an increase in premarital sexual intercourse, but the evidence is far from uniform. For instance Zaba report a consistent rise in age at first sex as well as a decline in premarital sex in Uganda, Ghana and Kenya (Zaba B et al., 2002). Mensch have also reported that age at first sex has remained the same or increased in many African countries (Mensch et al., 2006). Similarly, Blanc and Way find that in many sub-Saharan African countries both age at first sex and age at marriage have increased and that the gap between the timing of the two events has increased primarily due to a longer delay in first marriage than in first sex (Blanc and Way, 1998).

Changes in marriage pattern, such as delayed marriage, are believed to bring in the issues of dating, premarital sex, unwanted pregnancy, abortion, STDs and HIV/AIDS (Jones, 2007).Concern with special health needs of adolescents has also recently been growing in a world where young people are vulnerable to HIV/AIDS. Studies in parts of Kenya and Zambia show that teenage brides are contracting HIV at a faster rate than sexually active single girls in the same locales (Clarke, 2004). These relevant studies clearly indicate that an increase in age at marriage leads to a rise in premarital sex, and in the absence of contraception, in unwanted pregnancies and a rise in adolescent fertility (Jones, 2007). Thus, the effect of late marriage on reproduction through shortening the reproductive life span has been widely recognized.

Societies with later age at first marriage and first sexual intercourse have experienced decreased fertility rates while in traditional populations in Asia and Africa where age at first marriage and first sex is younger, high levels of fertility have been observed (Mensch et al., 2005). Delayed age at marriage directly affects completed fertility by reducing the number of years available for childbearing. The reason for decreased fertility rates is that delayed age at marriage and at sexual intercourse permits women to complete their education, build labor force skills, and develop career interests that compete with childbearing within marriage. These career interests in turn, motivate women to limit family size or widen the spacing of their children. In other words,

educational attainment, women's participation in the labor force and women's economic independence are factors associated with older age at marriage (Blossfeld, 1992).

As in most developing countries, early marriage and early sexual intercourse is prevalent in Ethiopia. The 2000 Ethiopian DHS further reveals that 31% of women aged 25-49 years were married before age 15 and 70% of women in the same age group were married before age 18. The median age at first marriage in the same year was 16 years (EDHS, 2000). The 2005 EDHS also maintains similar trends. The report shows that, nationwide, close to 34% women aged 25-49 were married before age 15 and 66% were married before age 18. In 2005, the median age at first marriage was 16.1years (EDHS, 2005). With regard to sexual intercourse in 2000 EDHS 30% of women aged 25-49 years have had sexual intercourse before age 15, 69 % of women in the same age group has had sexual intercourse before age 18. The median age at first intercourse for women age 25-49 was 16 years and it is identical to the median age at first marriage in the year (EDHS, 2000). In 2005 EDHS 32 % of women had sexual intercourse before age 15, 65 % of women has had sexual intercourse before age 18. The median age at first sexual intercourse for women age 25-49 years was 16.1 years (EDHS, 2005), which is identical to the median age at first marriage in the years. This suggests that Ethiopian women generally begin sexual intercourse at the time of their first marriage.

A study by the National Committee on Harmful Traditional Practices of Ethiopia (NCTPE) estimated the proportion married before the age of 15 are 57 percent. The same study shows that the practice occurs in its most extreme forms in northern Ethiopia, where girls are married as young as eight or nine years of age. Although age at early marriage is widely practiced in many parts of the country, rates in Amhara and Tigray region are much higher than the national average (82 percent in Amhara, 79 percent in Tigray, 64 percent in Benshangul, 64 percent in Gambella and 46 percent in Afar) (NCTPE, 2003). A study conducted in two woreda's of the Amhara region also shows that 14 percent of women were married before age of 10 years, 39 percent before age of 15 years, and 56 percent before age of 18 years (Population Council, 2004).

It is crucial to study the effect of various socio-economic and demographic factors which affect age at first marriage and first sexual intercourse in Rural Ethiopia. There are relatively few

empirical generalizations about the social and economic determinants of age at marriage and sexual intercourse. It is generally believed that rural tradition fosters early marriage, while urbanization and other forces of modernization lead to marriage postponement. Having these, this study examined factors associated age at first marriage and first sexual intercourse using survival models. Survival analysis is a statistical method for data analysis where the response variable is the time to the occurrence of an event, time-to-age at first marriage and time to age at first sexual intercourse in this study.

The study used Semi-parametric and parametric accelerated failure time (AFT) models fitted using weibull, log-logistic and log-normal baseline distributions to compare and get the best model which fits the time-to-age at first marriage and time to first sexual intercourse data appropriately.

1.3 Statement of the Problem

Age at marriage and age at first sexual intercourse are determinants of an array of population and health outcomes. Based on the United Nations (UN) Convention on the Rights of the Child, child marriage refers to marriage under age 18 (UN 2000). Marriage before the age of 18 is a fundamental human rights violation. Child marriage disproportionately affects young girls, who are much more likely to be married as children than young boys (Mathur et al 2003; UNICEF 2005; Save the Children 2004). The latest international estimates indicate that worldwide, more than 60 million women aged 20–24 were married before they reached the age of 18 (UNICEF 2007).

Early marriage and early sexual intercourse leads to early childbearing, with significantly higher maternal mortality and morbidity rates as well as higher infant mortality rates (Mensch 2005; UNICEF 2005; Save the Children2004). Moreover, child marriage has negative effects on girls' education. Girls with low levels of schooling are more likely to be married early, and child marriage virtually puts an end to a girl's education (Mathur et al 2003; Mensch 2005). A child bride's lack of education and peers limits her support systems, and without skills, mobility, and connections, she is constrained in her ability to overcome poverty for herself, her children, or her family. Young girls married to older men with more sexual experience are also at greater risk of

HIV infection (Clark 2004), and child brides are at heightened risk of violence in the home (Santhya et al 2010).

While most countries have placed the minimum age for marriage at 18 and the practice of child marriage has decreased globally over the last 30 years, it remains common in rural areas and very low-income areas (UNFPA, 2011). Impoverished parents often believe that child marriage will protect their daughters; however, it generally results in lost development opportunities, limited life options, and poor health.

The widespread execution of early family formation in most developing countries has attracted the attention of scholars from different disciplines, mainly because both the causes and the consequences of this practice are far reaching. Not only does early marriage pose serious health threats to girls who experience early sexual intercourse and childbearing following marriage formation; it also leads to social and economic underdevelopment both at individual and aggregate level (Clark, 2004; Assefa, *et. al*., 2005).

Ethiopia has extremely high rates of child marriage, particularly in rural areas. In fact, the Ethiopian Ministry of Health presents early marriage as one of the most prevalent types of harmful traditional practices that are performed in Ethiopia (Assefa, *et. al*., 2005). Even though, several measures have been taken by the governmental and Non-Governmental agencies in creating awareneous on age at first marriage the practice is available in rural areas. Most of the studies on age at first marriage and first sexual intercourse in Ethiopia were conducted by cross-sectional study and ordinary logistic regression model. This study has tried to fill the gap by explore factors that affect age at first marriage and age at first sexual intercourse in Rural Ethiopia based on the 2011 EDHS data using Cox proportional hazards model and parametric survival model. This study tries to address the following major research questions:

I. What are the key socio-economic and demographic predictors of age at first marriage and age at first sexual intercourse among women in Rural Ethiopia?
II. Which model is appropriate for analyzing the predictors of age at first marriage and age at first sexual intercourse?

1.4 Objective of the Study

1.4.1 General Objective

The General objective of the study is modeling age at first marriage and age at first sexual intercourse separately among women in Rural Ethiopia.

1.4.2 Specific Objectives

The specific objectives are:

I. To identify factors associated with age at first marriage and age at first sexual intercourse for women in Rural Ethiopia.
II. To compare the performance of Cox proportional hazards and accelerated failure time (AFT) model in modeling age at first marriage and age at first sexual intercourse.

1.5 Significance of the Study

The result of this study would provide information on age at first marriage and age at first sexual intercourse among women in Rural Ethiopia and their determinant factors. The results can provide an important input to design interventions to mitigate the problems of early marriage and early sexual intercourse.

2 Literature Review

2.1 Early Marriage

The term early marriage is used to refer to both formal marriages and informal unions in which a girl lives with a partner as if married before age of 18 (UNICEF 2005). Similarly, UNIFPA (2006) defines early marriage and child marriage, any marriage carried out below the age of 18 years, before the girl is physically, physiologically, and psychologically ready to shoulder the responsibilities of marriage and childbearing. Child marriage, on the other hand, involves either one or both spouses being children and may take place with or without formal registration, and under civil, religious or customary laws. Ango (1991) defined early marriage as either marriage of an adolescent girl/boy or marriage of an adolescent girl to a matured man or vice-versa and this happens at puberty when the individual is getting matured. The marital age according to Molokwu (2000) is above the age of 18 years, when the individual is physically, socially, academically and emotionally matured to cope with the challenges of marriage. Also, the UNICEF (2001) on the rights of the child recommends that children should not be separated from their parents before 18 years unless it is considered necessary.

2.1.1 Age at First Marriage

It is the time or age at which both male and female enter into marriage and their marital union is formed and assessed in years (United Nations, 2000). Demographers have focused on age at marriage because it signals exposure to the risk of pregnancy. Malthus (1872) advocated moral restraint or chaste postponement of marriage until one could provide for children as a means to balance population size with available resources. Comparatively late marriage did play an important role in reducing fertility levels in Europe (Hajnal, 1965), and age at marriage continues to be associated with fertility rates, especially in regions such as Africa where contraceptive use is uncommon (Adlakha and Kumar, 1991).

2.1.2 Risk Factors of Age at First Marriage

Age at First marriage is affected by a complex range of factors. Some of which are social, economical, religious and cultural norms. Studies elsewhere have, however, identified a number of factors that seem to influence the timing of marriage (Veronique, 2002). Some of the factors are region of residence, place of residence, education, occupation, economic status, religion and ethnicity. Increases in age at marriage are associated with major social-structural changes such as increases in educational attainment, urbanization, and the emergence of new roles for single women (United Nations, 1990). Jejeebhoy (1995) found that education is the single factor most strongly related to the postponement of marriage, but the relationship may be subject to threshold effects.

Shapiro (1996) has been argued that age at first marriage in rural areas was higher than in urban areas, mainly because of the higher average educational and employment status of women in urban areas, and the peer-level effects, which influences members of some urban communities to adopt a similar higher age of marriage as a norm behavior of pattern. Singh S. and Samara R. (1996) suggested, women in urban areas are exposed to modern values encouraging later marriage and are less likely to be under the influence of kin who control the timing of marriage and choice of spouse.

Kamal (2011) employed Cox proportional hazard regression analysis on Socio-economic determinants of age at first marriage of the Ethnic Tribal women in Bangladesh to the data set containing 792 observations from four different tribal communities compiled on the basis of a household survey. The mean of age at first marriage for females were found to be 18.9 years. The findings revealed that woman's educational attainment and pre-marital work status was significantly delayed the timing of marriage. Parents' economic status had the most significant effect on marital timing. The higher the economic status of the parents, the lower is the probability of age at earlier marriage. This is no doubt partly because parents with a low economic status are less likely to keep their daughters in school, due to financial difficulties.

In India, Johnson-Lans and Jones (2011) examines the relative importance of economic and social factors in determining the probability of a girl becoming a child bride. The findings reveal that economic factors is way behind social conformity in explaining child marriage as none of

economic variables such as household income, poverty status, and land ownership significantly affect the probability of child marriage in rural India. Qualitative study from Ghosh (2011) in West Bengal divides the causes of child marriage based on four different perspectives in a family; fathers, mothers, elders, and daughters. Fathers and elders are found to put poverty as the first order of preference for the cause of child marriage. Findings from qualitative studies on moving out of poverty in Indonesia also revealed that marriage is sometimes used as a way out of poverty (Febriany,2006). Mothers and daughters put perceptions that marriage is essential and lack of awareness respectively as the first order of reference. Sarkar (2009) uses logistic regression analysis in order to find determinants of early marriage in Bangladesh. His study shows that education, working status, husband's education, and places of resident exert the significant effect on early marriage. In case of Indonesia, Savitridina (1997) indicates that women's education, work status before marriage, husband's education, and current residence are the predictors for early marriage in Java, with education as the strongest one.

Joseph *et al.*(2012) using SUSENAS 2010 data as a nationally representative data for Indonesia in 2010 to examine the effect of socio-economic factors on Prevalence of Child Marriage and Its Determinants among Young Women in Indonesia by using Logistic Regression. The result showed that there were a negative correlation between age at child marriage and higher income of households, exposure to the media through the internet, education of household head, number of children in a family who are in high school and higher education, and access to the free healthcare. They also found that there were positive correlations between child marriage and the use of internet through cellular phone and number of children in a family who were in elementary school.

Bayisenga (2012) examined the issue of early marriage in sub-Saharan Africa. It sheds light specifically on reason behind its perpetuation, its harmful consequences which include barrier to education, enjoyment of girl’s human rights and how it further threatened the development of countries. He observed that early marriage is due to various factors including among others, the search for economic survival, protection of young girls, peer group and family pressure, controlling female behavior and sexuality, wars and civil conflicts and socio-cultural and religious values. It is a violation of girl’s human rights as it deprives her from freedom, opportunity for personal development, and other rights. It also a developmental challenge for

population pressure, health care costs and lost opportunities of human development. It is barrier to girl's education as young girls drop out from school to get married which impacts negatively on the community as whole and on the wellbeing of future generation. This practice stands in direct conflict with the objectives of the Millennium Development Goals (MDGs); such as the promotion of basic education, fight against poverty, the prevention of HIV/AIDS and reduction maternal mortality rate.

A recent study by UNICEF (2011) in six West African countries showed that 44 percent of women 20-24 years old in Niger were married under the age of 15. The needs to follow tradition, reinforce ties among or between communities, and protect girls from out-of-wedlock pregnancy were the main reasons given. In the communities studied, all decisions on the timing of marriage and the choice of spouse are made by the fathers. Type of place of residence may be equally important, as rural areas generally associated with age at early marriage. People living in urban areas are exposed to a more diverse lifestyle and subject to a weaker social control than those in rural areas. Rural areas tend to have institutional and normative structures such as the kinship and extended family that promote age at early marriage and childbearing (UN, 1990 and 2000). These social structures and networks were less potent and individual responsibility in the matters of marriage is emphasized in urban areas. Women in urban areas need to develop skills, gain resources, and achieve maturity to manage an independent household and thus they have to delay marriage.

Ikamari LD, (2005) investigated the effect of education on the timing of marriage in Kenya using the Cox proportional hazard model and linear regression analysis to analyze the data. The result showed that education has a statistically significant and strong positive effect on a woman's age at first marriage; the effect remaining robust in the presence of a number of controls. The highly educated women are more likely to delay marriage. Significant variations in the effect of education across the generations of women are apparent. The effect is greater for the younger women, indicating increased postponement of marriage. Premarital sexual activity, premarital childbearing, region of residence, religion and year of birth are also significantly associated with age at first marriage.

Determinants of child marriage among females are presumably believed as the factors behind female transition to marriage. Several researches are conducted in determining female age at first marriage. Lung Vu (2009) finds that education, place of residence, wealth, current age, region, and ethnicity are significantly related to age at first marriage in Vietnam while Agabaet al (2011) indicates that educational attainment, religion, district of residence, and birth cohort are strong socio-economic determinants of first marriage in Western Uganda. Both of these studies boldly underline the influence of educational attainment of girls in determining their age at first marriage. This finding is also supported by Manda and Meyer (2005) who find that women with higher levels of educational attainment are far more likely to enter marriage at later age than those without any or with little education, given of different model to approach. This confirms that the issue of development, in which the extended education for girls take place, provides a clearer explanation on why child marriage happens mostly in a lower educated girls characteristics (Jones, 2010; Singh and Samara, 1996).

Abdallah (2011) in a study observed that in Nigeria, the practice of child marriage is deeply entrenched in tradition, culture and religion and the country has one of the highest rates of child marriage in the world, with estimated 42 percent of girls married before 18 years; and while this is found among many ethnic groups across the country, its predominance is clearly in the northern part of the country. While nationwide, 20 percent of girls are married by age 15, and 40 percent are married by age 18, child marriage is extremely prevalent in some regions such as among the predominantly MuslimHausa-Fulani of the Northwest and North-East (of which Zaria is a part) where 48 percent of girls are married by age 15, and 78 percent are married by age 18.

Palamuleni (2011) researched on socioeconomic determinants of age at first marriage in Malawi. He used three approaches in his data analysis. Descriptive univariate analyses were used to inspect the frequency distributions of the various factors. Bivariate analysis was employed to examine the relationships of the independent variables and age at first marriage. Chi-square tests of independence were conducted for categorical variables. Logistic regression was used to examine the impact of social and economic factors on age at first marriage. He reports that out of the seven independent variables that were considered in the model, five of them were found to be the most important determinants of age at first marriage in Malawi and these include; age, region, education, religion and ethnicity. However, only age and education were statistically

significant. Rural-urban residence and wealth status were excluded from the model suggesting that these two variables were not significantly associated with the age at first marriage.

Adedokun, et. al., (2012) observed that child marriages are often arranged in two distinct ways, within a context of force and coercion, either by parents or other persons in positions of authority in the family arranging their young daughter's marriage to an adult, often a much older man or arranging the future marriage of two children. It is not uncommon to find girls of 7 -14 years already married off with the girls sent off to live with the families of the husbands. Generally, prospective husbands are selected based on social, religious and monetary factors and age is not considered an important factor, as husbands are on the average 12 years older than the child brides in monogamous unions and up to 15-20 years older in polygynous unions (Population Council, 2005) and in isolated cases, may be several decade solder.

Adebowale *et al*. (2012) used Cox proportional hazard models to determine Survival Analysis of Timing of First Marriage among Women of Reproductive age in Nigeria data on NDHS (2008). The result showed that the mean age at first marriage was 17.8±4.8 years. The covariates region, education, religion, and residence were significantly associated with the timing of first marriage ($p<0.001$). Hazard of age at early marriage also varied considerably among the subgroup of women in terms of their place of residence, education and religion. Women who reside in rural area (HR=1.15; CI=1.11-1.18) married earlier than their counterparts in the urban area. Age at first marriage was related to levels of education($p<0.001$). The data further showed that the higher the level of education of the women, the lower hazard of earlier marriage. Women who had primary, secondary and higher education were 0.88($p<0.001$), 0.68($p<0.001$) and 0.44($p<0.001$) respectively indicating less likely to marry earlier than those with no education. Muslim women married earlier (HR=1.34; CI=1.29 - 1.39) than Christians.

Studies on family transition that focus on specific regions in Ethiopia are major references on understanding the process of family formations and the high rates of early marriage prevalence in the country. In one recent paper, Erulkar (2013) studies early marriage and its relation to intimate partner violence in 7 regions of Ethiopia, namely, Amhara, Addis Ababa, Afar, Beneshangul-Gumuz, Oromia, SNNPR and Tigray by focusing on the factors associated in early marriage formation. Analysis of ever married women between 20 and 24 years old by age at first marriage

were made for a sample from population-based surveys that were conducted in 2009 and 2010, using multinomial logistic regression models. The results show that, the odds of having married before the age of 15 were 4 times higher for women living in rural areas compared to the ones living in urban areas. The odds were almost 3 times higher for rural-living women to be married between ages 15 and 17 compared to urban residing women. On the other hand, Muslim women had higher odds of having married between ages 15 and 17 (3.12) than before the age of 15 (1.76), compared to Orthodox Christian women and women with other religious affiliations. Moreover, women who have never attended school had 9 times higher odds of having married before the age 15, and 5 times higher to be married between 15 and 17 years old, compared to women who have ever attended school. Again, area and region of residence as well as education level are factors that can be attributed to both causes and consequences of early marriage.

However, the study also included qualitative survey results about the characteristics of adolescent marriages; whether the marriage was arranged, chosen, or through abduction; who was the main supporter of the marriage; whether the respondent knew the husband beforehand; and the reason for school drop-outs. The author have found that 71% of marriages between age 15 and 17, and 89% of marriages before age 15 were arranged by parents. In all types of marriages, the main supporter of the marriage was reported as mothers, moreover, as the age at marriage decreases, the likelihood of knowing about the marriage and the husband decrease and the age difference between spouses increase (Erulkar, 2013).

Tezera (2013) used data from 2011 EDHS to examine the effect demographic and Socio-economic variables to determine early marriage among women in Ethiopia using ordinary logistic regression and multilevel logistic regression model for analysis. The results of the ordinary logistic regression revealed that region, place of residence, religion of a woman, educational level of women, husband's/partner education level, respondents work status, households wealth index and exposure to mass media were found to be significant predictors for early marriage whereas occupation of husband's and number of siblings was not significant. Similarly from the category of religion of women protestant was insignificant at 5% level of significance.

2.2 Early Sexual Intercourse

Early sexual debut has been associated with risky sexual behavior and therefore analyzed for its relationship to HIV and other sexually transmitted diseases (Rosenberg, 2002). Early sexual debut also leads to early childbearing, which increases risks for poor health outcomes including obstructed labor and its attendant risks of maternal mortality and morbidity (Melah, 2007).

Age at first sex is also an important indicator of exposure to risk of pregnancy and sexually transmitted infections during adolescence. In fertility studies age at first marriage is often used as a proxy measure of the onset of a woman's exposure to pregnancy, but in many societies premarital sexual activity is common, and age at first sex would be a better proxy (Stover, 1998).The research literature identifies several factors behind the trend towards increased first sex outside of marriage and delayed marriage. These factors include increased access to health care and education, urbanization, and improved infant and child survival (Livi-Bacci M, 2012,). In addition to these factors which affect the demand for children, family planning communication and improvements in contraceptive supply also influence reproductive behavior (Bongaarts et al., 2012).

2.3 Review of Survival Models

The origin of survival analysis goes back to the time when life tables were introduced. Life tables are one of the oldest statistical techniques and are extensively used by medical statisticians and by actuaries. Yet relatively little has been written about their formal statistical theory. Kaplan and Meier (1958) gave a comprehensive review of earlier work and many new results. Cox (1972) was largely concerned with the extension of the results of Kaplan and Meier to the comparison of life tables and more generally to the incorporation of regression like arguments into life table analysis. Survival models have the capability of handling censored data. Cox, 1972 and Cox and Oakes, 1984 used survival analysis in modeling human lifetimes. Fergusson, 1984 used hazard functions to study the time to marital breakdown after the birth of child. Hazard functions had been also used in studies of time to shift in attentions in classroom Felmlee, 1983; in study of relapse of mental illness (Lavori, 1984), marital dissolutions Morgan, 1988 and human lifetimes (Gross et al., 1975).

Proportional hazards modeling is the most frequently used type of the survival analysis modeling in many research areas, having been applied to topics such as smoking relapse (Stevens& Hollis, 1989), affective disorders childhood family breakdown interruptions in conversation (Dress, 1986), and employee turnover (Morita *et al*., 1989), and in medical areas for identification of important covariates that have as significant impact on the response of the interested variables. Adebowale *et al*. (2012) used Cox proportional hazard models to determine Survival Analysis of Timing of First Marriage among Women of Reproductive age. Kamal (2011) employed Cox proportional hazard regression analysis on Socio-economic determinants of age at first marriage

3 Methodology

3.1 Data

3.1.1 Source of Data

The data for this study was extracted from the published reports of Ethiopian Demographic and Health Survey (EDHS, 2011). It is the third survey conducted in Ethiopia as part of the worldwide DHS project. The principal objective of the 2011 EDHS was to provide current and reliable data on marriage, fertility and family planning behavior, child mortality, adult and maternal mortality, children's nutritional status, use of maternal and child health services, knowledge of HIV/AIDS, and prevalence of HIV/AIDS and anemia.

3.1.2 Sampling Design

The sample for the 2011 EDHS was designed to provide population and health indicators at the national (urban and rural) and regional levels. The 2007 Population and Housing Census, conducted by the CSA, provided the sampling frame from which the 2011 EDHS sample was drawn. Administratively, regions in Ethiopia are divided into zones, and zones into administrative units called weredas. Each wereda is further subdivided into the lowest administrative unit, called kebele. During the 2007 census each kebele was subdivided into census enumeration areas (EAs), which were convenient for the implementation of the census.

The 2011 EDHS sample was selected using two stage cluster design and EAs were the sampling units for the first stage. The sample included 624 EAs, 187 in urban areas and 437 in rural areas. Households comprised the second stage of sampling. A complete listing of households was carried out in each of the 624 selected EAs from September 2010 through January 2011. A representative sample of 17,817 households was selected for the 2011 EDHS, of these, 16,702 were successfully interviewed. In the interviewed households 17,385 eligible women were identified for individual interview; complete interviews were conducted for 16,515. Women whose current ages are 15-49 years are included in the survey. Out of all 16,515 urban and rural respondents 10,417 rural women from nine regions and Dire Dewa city administration were included in the study. The data was analyzed using R and STATA statistical softwares.

3.1.3 Variables in the Study

The Response Variables

The response variables are age at first marriage and age at first sexual intercourse. Both are measured as the length of time from birth until the age at first marriage and age at first sexual intercourse which is measured in years. During the survey all women were asked a series of questions regarding to their marital status and sexual activity, whether they had ever lived with a man and they had sexual intercourse respectively. The response to these questions constitutes the women age at first marriage and age at first sexual intercourse and women who had not yet experienced the events resulting in right censoring of the data.

Explanatory Variables

Several predictors were considered in this study to investigate the determinant factors of age at first marriage and age at first sexual intercourse. These are women education level, region, religion, women work status, wealth index, and mass media. These covariates are described together with their coding scheme in table 3.1.All of these covariates are categorical.

Table 3. 1: Coding and explanation of explanatory variables

Variable	Description	Categories
Women education level	Women's level of education	0= No education;1= Primary; 2= Secondary,3=Higher
Region	Women's region	1=Tigray,2=Affar,3=Amhara, 4=Oromiya,5=Somali, 6=Benishangul-G, 7= SNNP, 8 =Gambela, 9 = Harari,10=Dire Dawa
Religion	Women's religion	1= Orthodox,2= Muslim, 3= Protestant,4= Others
Women work status	Women's Working status	0= Not had work,1= Had work
Wealth index	Household wealth index	1= Poor,2=Middle,3=Rich
Mass media	Access to mass media	0= No, 1= Yes

3.2 Method of Data Analysis

3.2.1 Survival Analysis

Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs. By time, we mean years, months, weeks, or days from the beginning of follow-up of an individual until an event occurs; alternatively, time can refer to the age of an individual when an event occurs. By event, we mean death, disease incidence, relapse from remission, recovery (e.g., return to work) or any designated experience of interest that may happen to an individual. The problem of analyzing time-to-event data arises in several applied fields such as medicine, biology, public health, epidemiology, engineering, economics, sociology, demography and etc. The terms lifetime analysis, duration analysis, event history analysis, failure-time analysis, reliability analysis, and transition analysis refer essentially to the same group of techniques although the emphases in certain modeling aspects could differ across disciplines (Aalen *et a*l., 2008).

The use of survival analysis, as different to the use of other statistical method, is most important when some subjects are lost to follow up or when the period of observation is finite certain patients may not experience the event of interest over the study period. In this latter case one cannot have complete information for such individuals. These incomplete observations are referred to as being censored. Most survival analyses consider a key analytical problem of censoring. In essence, censoring occurs when we have some information about individual survival time, but we do not know the survival time exactly. Such event can occur due to either; a person does not experience the event before the study ends, a person is lost to follow-up during the study period and, a person withdraws from the study for unknown/known reasons.
There are three categories of censoring.

i. Right censoring: Survival time is said to be right censored when it is recorded from its beginning to a defined time before its end time. This type of censoring is commonly recognized survival analysis and also considered in this study.
ii. Left censoring: Survival time is said to be left censored if an individual develops an event of interest prior to the beginning of the study.

Survival and Hazard Function

The survival time T is assumed to follow a distribution with density function f (t), then the survival function is given by $S(t) = P(T > t) = \int_t^{\infty} f(u)du$, and the hazard function $h(t) = f(t)/S(t)$. The relationship between the survival and the hazard function is given by: $S(t) = \exp\left(-\int_0^t h(u)du\right)$.

3.2.2 Non-parametric Survival Methods

Nonparametric methods are often very easy and simple to understand as compared to parametric methods. Furthermore, nonparametric analyses are more widely used in situations where there is doubt about the exact form of distribution. Survival data are conveniently summarized through estimates of the survival function and hazard function. The estimation of the survival distribution provides estimates of descriptive statistics such as the median survival time. These methods are said to be non-parametric methods since they require no assumptions about the distribution of survival time. Preliminary analysis of the data using non-parametric methods provides insight into the shape of the survival function for each group and get an idea of whether or not the groups are proportional, i.e., if the estimated survival functions for two groups are approximately parallel(do not cross). In order to compare the survival distribution of two or more groups, log-rank tests can be used.

The Kaplan-Meier (KM) estimator is the standard non parametric estimator of the survival function, S(t), proposed by Kaplan and Meier (1958) which is not based on the actual observed event and censoring times, but rather on the ordered in which events occur. It is also called the Product-Limit estimator. KM estimator incorporates information from all of the observations available, both censored and uncensored, by considering any point in time as a series of steps defined by the observed survival and censored times. When there is no censoring, the estimator is simply the sample proportion of observations with event times greater than t. The technique becomes a little more complicated but still manageable when censored times are included. Let ordered survival times are given by $0 \leq t_1 \leq t_2 \leq t_j \leq \infty$, then (Kaplan &Meier, 1958).

$$\widehat{S}(t) = \begin{cases} 1 & if\ t < t_1 \\ \prod_{j;t_j \le t}\left[1 - \frac{d_j}{r_j}\right] & if\ t \ge t_1 \end{cases} \quad (3.1)$$

Where; d_j is the observed number of events at time t_j and r_j is the number of individuals at risk at time t_j. The Kaplan-Meier estimator $\hat{S}(t)$ is a step function with jumps at the observed event times. The size of the jump at a certain event time tj depends on the number of events observed at t_j, as well as on the pattern of the censored event times before t_j.

3.2.3 Cox Proportional Hazards Regression Model

The Cox proportional hazards model is a very popular model in survival data. The Cox proportional hazards (PH) regression model (introduced in a seminar paper by Cox, 1972), a broadly applicable and the most widely used method of survival analysis.

Cox (1972) proposed a semi-parametric model for the hazard function that allows the addition of covariates, while keeping the baseline hazards unspecified and can take only positive values. With this parameterization the Cox hazard function is

$$h_i(t) = h_o(t)\exp\,(X_i^T\beta) \quad (3.2)$$

Where , $h_o(t)$ is the baseline hazard function which is obtained when all $X's$ are set to zero; X_i is a vector of covariates and β is a vector of parameters.

In this model, no distributional assumption is made for the survival time; the only assumption is that the hazards ratio does not change over time (i.e. proportional hazards model). Even though the baseline hazard is not specified, we can still get a good estimate for regression coefficients β, hazard ratio, and adjusted hazard curves.

The hazard ratio of two individuals with different covariates X and X^* is given by:

$$\widehat{HR} = \frac{h_o(t)\exp\,(\widehat{\beta}'X)}{h_o(t)\exp\,(\widehat{\beta}'X^*)} = exp\{\sum \widehat{\beta}'(X - X^*)\} \quad (3.3)$$

This hazard ratio is time-independent, that is why this is called the proportional hazards model. The parameter of the Cox proportional hazard model refers to the hazard ratio of one group in comparison to the other groups for categorical covariates and change in hazard ratio with a unit

change of the covariate for the continuous variables when other covariates are fixed. The change in hazard ratio for the continuous covariate is given by:

$$\frac{h_i(t,x_k+1\ )}{h_k(t,x_k\ )}=exp(\beta_k) \tag{3.4}$$

Which represent change in the hazard when there is a unit change in the covarite while other covariates keeps constant.

For catagorical explanatory variable X with a levels, the model containes (a-1) dummy variables defined as D_i = 1, if x =i, and 0 otherwise for i= 1,2,…,a -1. Let $\beta_1, \beta_2, \dots, \beta_{a-1}$ denote the coefficient of the levels of dummy variables. The ratio of the hazard of two subjects, one with X at level j and other with k (j,k = 1,2,…, a-1), provided that the value of all other explanatory variables for this subject are the same, the hazard ratio between these two categories is given by:

$$\frac{h(t\ |D_j\ )}{h(t\ |D_k\ )} = \frac{\exp\ (\beta_j)}{\exp\ (\beta_k)} = \exp(\beta_j - \beta_k) \tag{3.5}$$

The quantity $\exp(\beta_j - \beta_k)100\%$ signifies the ratio of hazard function for subject at level j and k of covariates, given that the effect of other covariate keeps fixed.

3.2.3.1 Model Diagnostics

After a model has been fitted, the adequacy of the fitted model needs to be assessed. The model checking procedures below are based on residuals.

Cox-Snell Residuals

The Cox -Snell residuals can be used to assess the overall fit of a model. The Cox–Snell residual is defined by:

$$rc_i = \exp(\beta^t X_i)\hat{\mathrm{H}}_{\mathrm{o}}(\mathrm{t_i}) = -\log\big(\hat{\mathrm{S}}(\mathrm{t_i})\big) \tag{3.6}$$

Where t_i is the observed survival time for individual i, X_i is the vector of covariate values for individual i, $\hat{H}_o(t_i)$ is the cumulative hazard function and $\hat{S}(t_i)$ is the estimated survival function on the fitted model.

Checking of Proportional Hazard Assumption

The validity of Cox's regression analysis relies heavily on the assumption of proportionality of the hazard rates of individuals with distinct values of a covariate. Proportional hazards means that the hazard function of one individual is proportional to the hazard function of the other individual, i.e. the hazard ratio is constant over time. There are several methods for verifying that a model satisfies the assumption of proportionality.

Graphical Method

We can obtain Cox PH survival function by the relation between hazard function and survival function

$$S(t,x) = S_o(t)^{\exp\left(\sum_{i=1}^{p}\beta_i x_i\right)} \qquad (3.7)$$

Where $x_1 = (x_1, x_2, \dots, x_p)'$ is the values of the vector of explanatory variables for a particualr individual. When taking the logarithm twice, we can easily get

Log[-log S(t,x)]=$\sum_{i=1}^{p}\beta_i x_i + \log\left[-\log S_o(t)\right]$.

Then the difference in log-log curves corresponding to two different individuals with variables

$x_1 = (x_{11}, x_{12}, \dots, x_{1p})$ and $x_2 = (x_{21}, x_{22}, \dots, x_{2p})$ is given by

Log[-log $S(t,x_1)$]−log[-log $S(t,x_2)$] = $\sum_{i=1}^{p}\beta_i(x_{1i} - x_{2i})$, which does not depend on t. this relationship is very helpful to help us identify situations where we may have proportional hazards. By plotting estimated log(-log(survival)) versus survival time for two groups we would see parallel curves if the hazards are proportional. This methods does not work well for continuous predictors or categorical predictors that have many levels because the graph becomes "cluttered". Furthermore, the curves are sparse when there are few time points and it may be difficult to tell how close to parallel is close enouph.

However, looking at the K-M curves and log(-log(survival)) is not enough to be certain of proportionality since they are univariate analysis and do not show whether hazards will still be proportional when a model includes many other predictors. But they support our argument for proportionality. We will show some other statistical methods for checking the proportionality.

Adding Time-dependent Covariates in the Cox Model

We create time-dependent covariates by creating interactions of the predictors and a function of survival time and including them in the model. For example, if the predictor of interest is X_j , then we create a time-dependent covariate $X_j(j) = X_j * g(t)$; where $g(t)$ is a function of time, example, t, log t or Heaviside function of t. The model assessing PH assumption for $X_j$adjusted for other covariates is

$$h\big(t, x(t)\big) = h_o(t)\exp\left(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_j x_j + \cdots + \beta_p x_p + \delta x_j * g(t)\right) \qquad (3.8)$$

here $x(t) = (x_1, x_2, \ldots, x_j(t))$ is the value of the vector of explanatory variables. The null hypothesis to check proportionality is that $\delta = 0$. The test statistic can be carried out using either a Wald test or a likelihood ratio test. In the Wald test, the test statistic is $W = \left(\frac{\hat{\delta}}{se(\hat{\delta})}\right)^2$. The likelihood ratio test calculates the likelihood under null hypothesis, L_o and the likelihood under the alternative hypothesis, L_a. The LR statistic is then $LR = -2\ln\left({L_o}/{L_a}\right) = -2(L_o - L_a)$, where L_o, L_a are log likelihood under two hypothesis respectively. Both statistics have a chi-square distribution with one degree of freedom under the null hypothesis. If the time-dependent covariate is significant, i.e. the null hypothesis is rejected, and then the predictor is not proportional. In the same way, we can also assess the PH assumption for several predictors simultaneously.

3.2.4 Accelerated Failure Time Model

Although parametric models are very applicable to analyze survival data, there are relatively few probability distributions for the survival time that can be used with these models. In these situations, the accelerated failure time model (AFT) is an alternative to the PH model for the analysis of survival time data. Under AFT models we measured the direct effect of the explanatory variables on the survival time instead of hazard. This characteristic allows for an easier interpretation of the results because the parameters measure the effect of the correspondent covariate on the mean survival time.

The AFT model states that the survival function of an individual with covariate X at time t is the same as the survival function of an individual with a baseline survival function at a time $t^{*}exp(a'X)$, where $a'= (a_1, a_2, \ldots, a_p)$ is a vector of regression coefficients. In other words, the accelerated failure-time model is defined by the relationship.

$$S(t|X) = S_o\{t * \exp(a'X)\}, \text{ for all X} \quad (3.9)$$

The corresponding log-linear form of the AFT model with respect to time is given analogous to the classical linear regression approach. In this approach, the natural logarithm of the survival time Y =log (T) is modeled. This is the natural transformation made in linear models to convert positive variables to observations on the entire real line. A linear model is assumed for Y;

Y = log(T)= $\mu + a'x + \sigma\epsilon$, where μ is intercept, $a' = (a_1, a_2, .., a_p)$ is a vector of regression coefficients , σ is scale parameter and ϵ is the error distribution assumed to have a particular distribution.

When we denote by S_0 the survival function when X = 0 then we find that

$p(T > t|X) = p(Y > \log(t)\ |X\ )$

$= p(\ \mu + \sigma\epsilon > \log(t) - a'X\ |X)$

$= p\{\exp\ (\mu + \sigma\epsilon) > t * \exp\ (-a'X)\ |X\}$

$= S_o\{t * \exp(-a'X)\}$

The effect of the covariates on the survival function is that the time scale is changed by a factor exp $(-a'X)$, and we call this an acceleration factor.

We note that when

exp $(-a'X) > 1 \rightarrow$ the survival process accelerates.

exp $(-a'X) < 1 \rightarrow$ the survival process decelerates.

To use an AFT model, a distribution must have a parameterization that includes a scale parameter. The logarithm of the scale parameter is then modeled as a linear function of the covariates. In this study the Weibull AFT, log- logistic AFT, and log-normal AFT Models were considered.

a. The Weibull AFT Model

The Weibull distribution is very flexible model for time-to-event data. It has a hazard rate which is monotone increasing, decreasing, or constant. It is the only parametric regression model which

has both a proportional hazards representation and an accelerated failure-time representation. It is a two-parameter model ($\lambda \, and \, \rho$), where λ is the scale parameter and ρ is the shape parameter because it determines whether the hazard is increasing, decreasing, or constant over time i.e., the hazard rate increases when, $\rho > 1$ and decreases when $\rho < 1$ as time goes on. When $\rho = 1$, the hazard rate remains constant, which is the special case of exponential distribution.

The pdf is given by;

$$f(t) = \lambda\rho t^{\rho-1}\exp(-\lambda t^{\rho}), where; \lambda, \rho > 0 \quad (3.10)$$

And the corresponding survival function and hazard for weibull distribution are given as;

$$S(t) = \exp(-\lambda t^{\rho}) \qquad h(t) = \lambda\rho t^{\rho-1} \quad (3.11)$$

From equation (3.7), The AFT representation of the survival and hazard function of the Weibull model is given by:

$$S_{ei} = \exp(-\exp\left(\frac{logt-(\mu+a'x)}{\sigma}\right) = \exp(-\exp(\frac{-(\mu+a'x)}{\sigma} t^{1/\sigma})) \quad (3.12)$$

$$h_i(t) = \frac{1}{\sigma} t^{1/\sigma-1}\exp(\frac{-\mu-a'x}{\sigma}) \quad (3.13)$$

b. Log-logistic AFT Model

The log-logistic distribution has a fairly flexible functional form, it is one of the parametric survival time models in which the hazard rate may be decreasing, increasing, as well as hump shaped that is it initially increases and then decreases. In cases where one comes across to censored data, using log-logistic distribution is mathematically more advantageous than other distributions. The log-logistic distribution is very similar in shape to the log-normal distribution, but is more suitable for use in the analysis of survival data. The log-logistic model has two parameters λ and ρ, where λ is the scale parameter and ρ is the shape parameter.

Its pdf is given by;

$$f(t) = \frac{\lambda\rho t^{\rho-1}}{(1+\lambda t^{\rho})^2} \quad (3.14)$$

The corresponding survival and hazard functions are given by;

$$S(t) = \frac{1}{1+\lambda t^{\rho}} \qquad h(t) = \frac{\lambda \rho t^{\rho-1}}{1+\lambda t^{\rho}} \tag{3.15}$$

Where, $\lambda \in R, \rho > 0$

When $\rho \leq 1$, the hazard rate decreases monotonically and when $\rho > 1$, it increases from zero to its maximum point and then decreases to zero. Suppose that the survival times have log-logistic distribution with parameter λ and , under the AFT model, the hazard function for the i[th] individual is

$$h_i(t/x) = h_o(texp(-a'x_i))\exp(-a'x_i) = \frac{\rho exp\left((\lambda)texp(-a'x_i\right)}{1+\exp(\lambda)\{texp(-a'x_i)\}^{\rho}} \tag{3.16}$$

The log-logistic AFT model with a covariate x may be written as;

Y= logT=$\mu + a'x + \sigma\epsilon$, where, $a' = (a_1, a_2, .., a_p)$, ϵ has standard logistic distribution. The survival and hazard function with covariate x is given as follows:

$$S_T(t/x) = \frac{1}{1+\lambda\exp(\beta'x)t^{\rho}} = \frac{1}{1+\exp(log\lambda+\beta'x)} \tag{3.17}$$

$$h_T(t/x) = \frac{\rho t^{\rho-1}\lambda\exp(a'x)}{1+\lambda\exp(a'x)t^{\rho}} = \frac{\rho t^{\rho-1}\lambda\exp(a'x)}{1+\exp(log\lambda+a'x)} \tag{3.18}$$

To interpret the factor exp $(\beta'x)$ for log-logistic model, one can notice that the odds of survival beyond time t for log-logistic model are given by $\frac{S_T(t)}{1-S_T(t)}$.

We can see that the log-logistic distribution has the proportional odds (PO) property. So this model is also a proportional odds model, in which the odds of an individual surviving beyond time t are expressed as:

$$\frac{S_T(t)}{1-S_T(t)} = \exp(-a'x)\frac{S_o(t)}{1-S_o(t)} \tag{3.19}$$

The factor exp $(-a'x)$ is an estimate of how much the baseline odds of survival at any time changes when individual has covariate x. And exp $(-a'x)$ is the relative odds of experiencing the event for an individual with covariate x relative to an individual with the baseline characteristics. As this representation of log-logistic regression is as accelerated failure time model with a log-logistic baseline survival function, then the log logistic model is the only parametric model with both a proportional odds and an accelerated failure-time representation. If T_i has a log-logistic distribution, then ε_i has a logistic distribution. The survival function of logistic distribution is given by (Collett, 2003)

$$S_{\varepsilon_i}(\varepsilon) = \frac{1}{1+\exp(\varepsilon)} \quad (3.20)$$

Then, the AFT representation of log-logistic survival function is given by

$$S_t(t) = \left[1 + t^{\frac{1}{\sigma}} \exp\left(\frac{-\mu - a'x}{\sigma}\right)\right]^{-1} \quad (3.21)$$

And the associated hazard function for the ith individual is given by

$$h_t(t) = \frac{1}{\sigma t}\left[1 + t^{\frac{-1}{\sigma}} \exp\left(\frac{-\mu - a'x}{\sigma}\right)\right]^{-1} \quad (3.22)$$

c. Log-normal AFT Model

If the survival times are assumed to have a log-normal distribution, the baseline survival function and hazard function are given by:

$$S_O(t) = 1 - \Phi\left(\frac{logt-\mu}{\sigma}\right) \quad (3.23)$$

$$h_o(t) = \frac{\phi\left(\frac{logt}{\sigma}\right)}{\left[1-\Phi\left(\frac{logt}{\sigma}\right)\right]\sigma t} \quad (3.24)$$

Where μ and $\sigma$are parameters, $\phi(x)$ is the probability density function and $\Phi(x)$ is the cumulative density function of the standard distribution. The survival function for the ith individual is

$$S_i(t) = S_o(t * \exp(\mu + a'x_i)) = 1 - \Phi\,\frac{(logt - a'x_i - \mu)}{\sigma} \quad (3.25)$$

Therefore the log survival time for the ith individual has normal$(\mu + a'x_i, \sigma)$. The log normal distribution has the AFT property. In a two group study, we can easily get

$\phi^{-1}[1 - S(t)] = \frac{1}{\phi}(logt - a'x_i - \mu)$, where x_i is the value of a categorical variable which takes the value 0 in one group and 1 in the other group. This implies that the plot $\phi^{-1}[1 - S(t)]$ against log(t) will be linear if the log-normal distribution is appropriate.

3.2.4.1 Parameterization

When we say proportional hazards (PH) it means that the hazard function of a group is proportional to the hazard function of the other group, i.e., the hazard ratio is constant over time (Klein, 1992). The hazard ratio is hence given by; HR = exp($\beta'X_{ij}$), where $\beta' = (\beta_1, \beta_2, \ldots..\beta_p)$ is a vector of regression coefficients and X_{ij} is the vector of covariates for subject j in cluster i. On the other hand, the accelerated failure-time (AFT) model describes stretching out or contraction

of survival time as a function of predictor variables. The acceleration factor which is usually denoted by ∅ is given by $\exp(\alpha' X_{ij})$ where $\alpha' = (\alpha_1, \alpha_2 \ldots \alpha_p)$ is a vector of regression coefficients in case of AFT model. For the exponential, weibull and log logistic survival model, the relationship between α and β is given by (Dätwyler and Stucki, 2009)

a. For exponential $\beta_j = -\alpha_j$, the exponential PH and AFT are in fact the same model, except that the parameterization is different, and hence HR=exp $(-\alpha_j)$ is the hazard ratio of the j^{th} group with the reference groups.

b. For weibull, $\beta_j = -\alpha_j\rho$, where ρ is the shape parameter and hence, HR=exp $(-\alpha_j\rho)$ is the hazard ratio of the j^{th} group with the reference groups.

c. For log-logistic $\beta_j = -\alpha_j\rho$, where ρ is the shape parameter and OR = exp$(-\alpha_j\rho)$ indicates the failure odds ratio of the jth group with the reference groups. The log-logistic model is a proportional odds (PO) model, i.e. it has constant OR for two groups.

3.2.4.2 Parameter Estimation of AFT Model

Parameters of AFT models can be estimated by maximum likelihood method. The likelihood of n observed survival times, $t_1, t_2, t_3 \ldots, t_n$, the likelihood function for right censored data is given by:

$$L(\alpha, \mu, \sigma) = \prod_{j=1}^{n} f_i(t_i)^{\delta_i} * S_i(t_i)^{(1-\delta_i)} \quad (3.26)$$

Where $f_i(t_i)$ is the density function of the i^{th} individual at time t_i, $S_i(t_i)$ is the survival function of the i^{th} individual at time t_i, δ_i is indicator variable.

3.2.4.3 Model Selection

For comparing models that are not nested, the Akaike's information criterion (AIC) is used which is defined as:

$$AIC = -2LogL + 2(k+c+1) \quad (3.27)$$

Where k is the number of covariates and c the number of model specific distributional parameters. Lower values of the AIC suggest a better model.

3.2.4.4 Model Diagnostics: AFT Model

Checking the Adequacy of Parametric Baselines

The graphical methods can be used to check if a parametric distribution fits the observed data. Model with the weibull baseline has a property that the log (-log(S(t)) is linear with the log of time, where S(t) = exp(-λt^{ρ}). Hence, log(-log(S(t))) = log(λ) + ρlog(t). This property allows a graphical evaluation of the appropriateness of a Weibull model by plotting log(–log($\hat{S}(t)$)) versus log(t) where $\hat{S}(t)$ is Kaplan-Meier survival estimate (Dätwyler and Stucki, 2009). The log-failure odd versus log time of the log-logistic model is linear. Where the failure odds of log logistic survival model can be computed as:

$$\frac{1-S(t)}{S(t)} = \frac{\frac{\lambda t^{\rho}}{1+\lambda t^{\rho}}}{\frac{1}{1+\lambda t^{\rho}}} = \lambda t^{\rho} \qquad (3.28)$$

Therefore, the log-failure odds can be written as:

$$\text{Log}\left(\frac{1-S(t)}{S(t)}\right) = \log(\lambda t^{\rho}) = \log(\lambda) + \rho\log(t) \qquad (3.29)$$

The log-logistic baseline can be graphically evaluated by plotting log ($\frac{1-\hat{S}(t)}{\hat{S}(t)}$) versus log(time), where $\hat{S}(t)$ is Kaplan-Meier survival estimate (Dätwyler and Stucki, 2009). If the plot is straight line, log-logistic distribution fitted the given data well.

For the log-normal distribution the plot $\Phi^{-1}[1 - \hat{S}(t))]$ versus log (t) should be linear.

The Residual Plots

For the parametric regression problem, analogs of the semi parametric residual plots can be made with a redefinition of the various residuals to incorporate the parametric form of the baseline hazard rates (Klein, 2012). The first such residual is the Cox–Snell residual that provides a check of the overall fit of the model. The Cox–Snell residual, r_j, is defined by:

$$r_j = \hat{H}(T_j \mid X_j) \qquad (3.30)$$

where $\hat{H}$ is the cumulative hazard function of the fitted model. If the model fits the data, then the r_j's should have a standard (λ =1) exponential distribution, so that a hazard plot of r_j

versus the Nelson–Aalen estimator of the cumulative hazard of the r_j's should be a straight line with slope 1.

The Quantile -Quantile Plot

A quantile-quantile or q-q plot is made to check if the accelerated failure time model provides an adequate fit to the data. The plot is based on the fact that, for the accelerated failure-time model,

$$S_1(t) = S_0(Ø\ t) \qquad (3.31)$$

Where S_0 and S_1 are the survival functions in the two groups and Ø is the acceleration factor. Let t_{op} and t_{1p} be the p^{th} percentiles of groups 0 and 1, respectively, that is:

$$t_{kp}=s_k^{-1}(1-p),\ k=0,1.$$

Using the relation $S_1(t) = S_0(Ø\ t)$, we must have $s_o(t_{op}) = 1-p = s_1(t_{1p}) = s_o(Øt_{1p})$for all t. If the accelerated failure time model holds, $t_{op} = Øt_{1p.}$ To check this assumption we compute the Kaplan–Meier estimators of the two groups and estimate the percentiles t_{1p}, t_{0p}, for various values of p. If we plot the estimated percentile in group 0 versus the estimated percentile in group 1 (i.e., plot the points t_{1p}, t_{0p} for various values of p), the graph should be a straight line through the origin, if the accelerated failure time model holds. If the curve is linear, a crude estimate of the acceleration factor q is given by the slope of the line (Klein, 1992).

4 Results

4.1 Exploratory Data Analysis

A total of 10,417 rural women from nine regions and Dire Dawa city administration were included in the study. Demographic and socio economic characteristics for the study respondents are summarized in table 4.1.The majority (64.3%) of these respondents had no education and comes from poor households (53.7%) according to wealth index classification. More than half (64.6%) of respondents had no any access to mass media and 70% of them had no work at the time of the survey.

Table 4. 1: Descriptive Summary of Demographic and Socio-economic Variables

Variable	Categories	Frequency	Percent
Region	Tigray	1269	12.2
	Affar	1025	9.8
	Amahara	1689	16.2
	Oromiya	1673	16.1
	Somali	575	5.5
	Benshangul-G	1034	9.9
	SNNP	1570	15.1
	Gambella	800	7.7
	Harari	423	4.1
	Dire Dewa	359	3.4
Women education level	No education	6701	64.3
	Primary	3425	32.9
	Secondary	210	2
	Higher	81	0.8
Mass media	No	6731	64.6
	Yes	3686	35.4
Wealth index	Poor	5589	53.7
	Middle	2062	19.8

	Rich	27666	26.6
Women work status	No	7289	70
	Yes	3128	30
Religion	Orthodox	3655	35.1
	Muslim	4406	42.3
	Protestant	2041	19.6
	Others	315	3

In this study we were interested in understanding the two important transitions women's make during their life time. These are age at first marriage (AFM) and age at first sexual intercourse (AFSI). As explained in the previous chapter, we have a semi-current status data for the two responses. From each respondent we have age at the time of the survey, information on whether the respondent had a marriage and sexual intercourse experience before, if so at what age the respondents engaged in to marriage? Respondents who had not married is considered as right censored, and their age at the time of the survey were used as a right censoring time. A Respondent who were engaged in marriage before but unable to recall the age at first marriage were considered as left censored and the age at the time of the survey is used as left censoring time. The same definition is also used for the time to first sex analysis.

Out of 10,417 women interviewed, 2,145(20.6%) did not get married at the time of the survey and none of the respondents had a marriage experience but unable to recall the age at first marriage (Left censored). Majority of the respondants, 8,272(79.4%) were married at the time of the survey or had an experience before. The overall median survival time of age at first marriage for Rural Ethiopian women is 17 years (with 95% CI; 16.90, 17.10).

With regard to their sexual intercourse out of 10,417 women interviewed, 2,069(19.9%) did not had sexual intercourse at the time of the survey and none of the respondents had this experience but unable to recall the age at first sexual intercourse also (Left censored). Majority of the respondants, 8,348(80.1%) have had sexual intercourse at the time of the survey or had an experience before. From women who had sexual intercourse at the time of the survey 61.4% of them had sexual intercourse at their first union. This suggested that the majority of Rural Ethiopian women begin sexual intercourse at the time of their first marriage. The rest (18.7%) of

the women enter into sexual intercourse at the age of 8-35 years. The overall median survival time of age at first sexual intercourse for Rural Ethiopian women is 17 years (with 95% CI; 16.90, 17.11), these is similar to the median age of first marriage.

The analysis begins by graphical exploration of the data to visualize the distribution of the survival times with respect to the recorded variables for the two responses. Figure 4.1 and 4.2 below illustrates the overall survival curve for all subjects. From the Kaplan Meier Curve, it can be concluded that most of the rural women engaged in to marriage and sexual intercourse just after their age turn in to a double digit (in their 10^{th} to 15^{th} years). Similar plots are also produced by baseline characteristics of the respondent (Appendix C). We also perform non-parametric test and the result is presented in table 4.2.

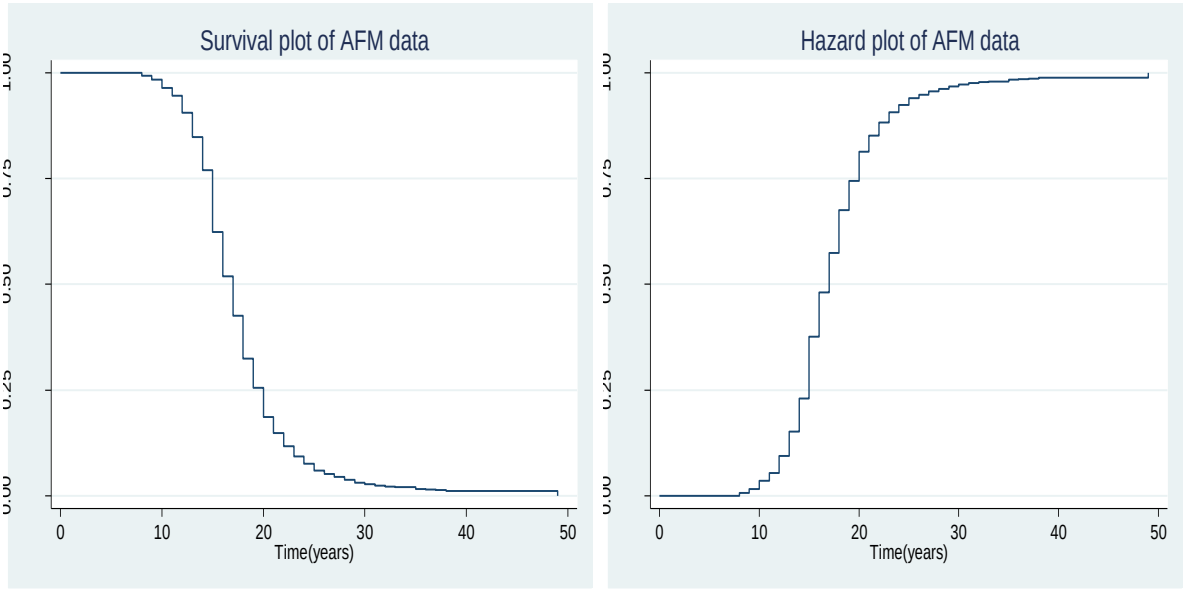


Figure 4. 1: The K-M plots of Survival and hazard function of age at first marriage

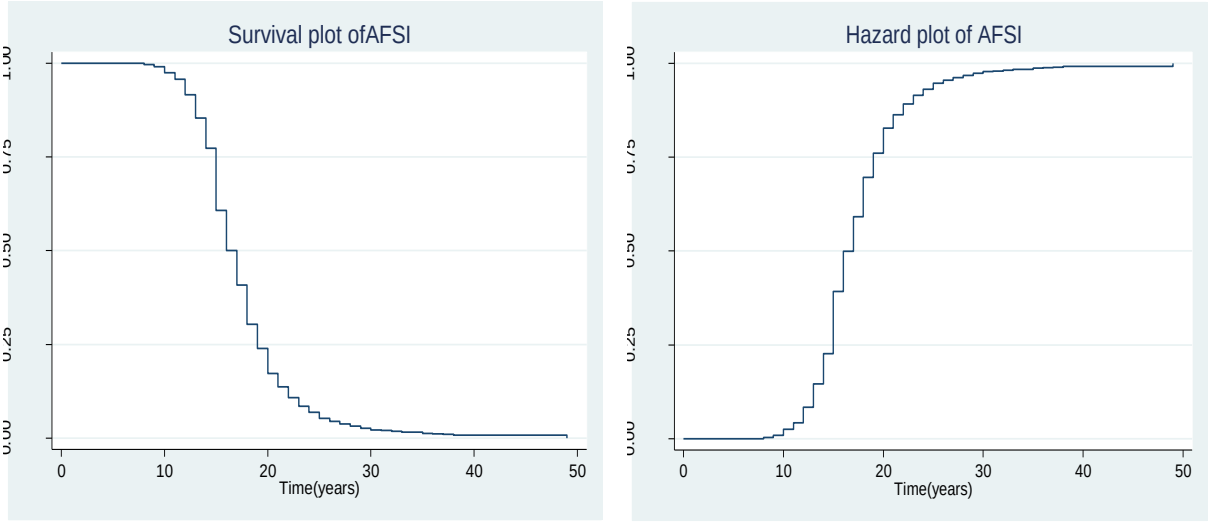


Figure 4. 2: The K-M plots of Survival and hazard function of age at first sexual intercourse

The Kaplan-Meier survivor estimator is used to investigate the significance differences between the survival probabilities of different variables. The survival plot for age at first marriage and sexual intercourse by region of women in Rural Ethiopia (Appendix C) show that women from Dire Dawa was married and enter into sexual intercourse later when compared with others and women from Amhara region had less survival than women from other regions. The result of the log rank test table 4.2 also showed that the difference is statistically significant at 5% level of significance (p= 0.000). In order to investigate the significant difference between the survival of women by women's educational level, Kaplan-Meier survivor estimates for four education level were plotted in Appendix C. The differences that are displayed in survival curve emphasize that women who had higher education was married and had sexual intercourse later when compared with others until the median time, and women who had no education had less survival than educated. As we indicated from the graph there is a high gap at the mid time between a primary and secondary educated women on marriage and sexual intercourse. This shows that women who had secondary education were more survived than uneducated and primary educated women. The log rank test table 4.2 also support the difference is statistically significant at 5% level of significance (p=0.000).

The survival plot of age at first marriage and first sexual intercourse by access to mass media (in Appendix C), as it can be observed from the plot, the survival curve for both groups is overlapped from the beginning to the end. This implied that age at first marriage and first sexual

intercourse for women who had access to mass media and who didn't have access is the same. The result of the log rank test (table 4.2) is not significant (p= 0.174) and (p=0.490) for first marriage and first sexual intercourse respectively, these indicates that the difference is not statistically significant at 5% level of significance. The KM curves for wealth index given in Appendix C. The plot suggested that the risk of marriage and sexual inter course is similar for all groups (poor, middle, and rich) at the beginning of the plot. But a little difference is observed at the middle of the curve. At the middle point of the curve, the survival plot age at first marriage for poor income family is below that of the middle and rich family. This indicated that women from poor income family are married and had sexual intercourse before those from middle income and rich family. The log rank test also significant (table 4.2) this means that the difference is statistically significant at 5% level of significance.

With regard to the survival plot of age at first marriage and first sexual intercourse by women's work status (in Appendix C).The survival plot of age at first marriage and sexual intercourse for women who had not work is below that of women who had work. This implied that women who had not work married and had sexual intercourse earlier than those who had work. This shows that those women who had work are prolonged age at first marriage and first sexual intercourse. The result of the log rank test in table 4.2 also support the existence of significant difference (p= 0.00124 and p=0.0113). The survival plot of age at first marriage and first sexual intercourse by religion (in Appendix C). Women who follow Protestant religion were prolonged age at first marriage and first sexual intercourse relatively. The result of the log rank test (table 4.2) also significant (p=0.000) means this difference is statistically significant at 5% level of significance.

Table 4. 2: Median Time of First Marriage and First Sexual Intercourse result of Non-parametric test

Variables	Age at First Marriage			Age at First Sexual Intercourse		
		Log-rank test			Log-rank test	
	Median (95% CI)	Chisq	P-value	Median (95% CI)	Chisq	p-value
Region		581	0.000		600	0.000
Tigray	17(16.71,17.28)			16(15.80,16.20)		
Affar	16(15.73,16.26)			16(15.74,16.26)		
Amhara	15(14.85,15.14)			15(14.87,15.13)		
Oromiya	17(16.76,17.23)			17(16.77,17.23)		
Somali	17(16.68,17.31)			17(16.69,17.30)		
Benishangul-G	16(15.77,16.23)			16(15.78,16.23)		
SNNP	18(17.29,18.27)			18(17.74,18.26)		
Gambela	16(15.67,16.32)			16(15.81,16.19)		
Harari	16(15.56,16.43)			16(15.54,16.46)		
Dire Dewa	18(17.45,18.54)			18(17.59,18.40)		
Women education Level		964	0.000		901	0.000
No education	16(15.91,16.08)			16(15.924,16.076)		
Primary	18(17.80,18.2)			18(17.808,18.192)		
Secondary	24(21.17,26.82)			22(20.695,23.305)		
Higher	22(20.92,23.07)			20(19.042,20.958)		
Mass media		1.8	0.174		0.5	0.490
No	17(16.87,17.12)					
Yes	17(16.83,17.16)			16(15.87,16.13)		
				17(16.84,17.16)		
Wealth index		35.7	1.78e-08		30.7	2.13e-07
Poor	17(16.87,17.13)			16(15.87,16.13)		
Middle	17(16.77,17.22)			17(16.78,17.22)		
Rich	17(16.79,17.20)			17(16.79,17.20)		
Women work Status		9.8	0.00177		6.4	0.0113
Not had work	17(16.88,17.11)			16(15.886,16.114)		
had work	17(16.79,17.20)			17(16.805,17.195)		
Religion		136	0.000		180	0.00
Orthodox	16(15.81,16.19)			16(15.869,16.131)		
Muslim	17(16.86,17.13)			17(16.861,17.139)		
Protestant	18(17.77,18.22)			18(17.778,18.222)		
Others	17(16.45,17.54]			17(16.482,17.518)		

4.2 Cox Proportional Hazards Model

Model fitting commenced with univariable models on the training set. This was important in not only identifying factors to be included in the multivariable model, but also in identifying relevant predictive factors influencing age at first marriage and sexual intercourse independent of the other factors. Any predictor having significant association with these outcomes at the 20% level in univariable models was included in the multivariable model. In multivariable analysis, statistical significance was attained if P<0.05. All P values were two-tailed. Table 4.3 below presents the results from the univariable and multivariable analysis for the two responses. In univariable analysis of Cox proportional hazards model, all variables are significantly associated with age first marriage at 20% level but for age at first sexual intercourse access to mass media was insignificant at 20% level. Therefore those variables that are significant in univariable analysis were included in the final model (multivariable analysis).

Table 4. 3: Univariable Analysis of AFM and AFSI using Cox PH Model

Variables	Age at First Marriage			Age at First Sexual Intercourse		
	Coef	HR (95%CI)	p-value	Coef	HR (95%CI)	p-value
Region						
Tigray	Ref					
Affar	0.09	1.09(0.997,1.196)	0.056	-0.1	0.9(0.826,0.989)	0.029
Amhara	0.44	1.55(1.423, 1.677)	0.000	0.24	1.27(1.171,1.378)	0.000
Oromiya	-0.15	0.86(0.795, 0.939)	0.001	-0.32	0.72(0.666,0.786)	0.000
Somali	0.01	1.01(0.907, 1.125)	0.854	-0.18	0.83(0.749,0.929)	0.001
Benishangul-G	0.21	1.23(1.123, 1.348)	0.000	0.02	1.02(0.928,1.114)	0.714
SNNP	-0.4	0.67(0.613, 0.728)	0.000	-0.59	0.56(0.51,0.606)	0.000
Gambela	0.1	1.1(1.000,1.218)	0.049	0.05	1.06(0.958,1.162)	0.274
Harari	0.11	1.12(0.988, 1.262)	0.078	-0.09	0.92(0.811,1.035)	0.159
Dire Dewa	-0.29	0.75(0.656, 0.852)	0.000	-0.49	0.61(0.539,0.701)	0.000
Women education Level						
No education	Ref					
Primary	-0.66	0.52(0.492, 0.545)	0.000	-0.65	0.53(0.499,0.55)	0.000
Secondary	-1.63	0.2(0.157, 0.245)	0.000	-1.5	0.22(0.179,0.27)	0.000
Higher	-1.43	0.24(0.177, 0.322)	0.000	-1.08	0.34(0.261,0.444)	0.000
Mass media						
No	Ref					

Yes	-0.03	0.97(0.926,1.014)	0.171	-0.02	0.98(0.941,1.029)	0.484
Wealth index						
Poor	Ref					
Middle	-0.06	0.94(0.893,0.999)	0.047	-0.07	0.93(0.882,0.988)	0.017
Rich	-0.16	0.85(0.807,0.896)	0.000	-0.15	0.86(0.819,0.908)	0.000
Women work						
Status						
Not had work	Ref					
Had work	-0.08	0.92(0.882,0.969)	0.001	-0.06	0.94(0.895,0.983)	0.007
Religion						
Orthodox	Ref					
Muslim	-0.05	0.95(0.903,0.996)	0.033	-0.14	0.87(0.830,0.915)	0.000
Protestant	-0.34	0.71(0.667,0.756)	0.000	-0.41	0.67(0.626,0.709)	0.000
Others	-0.14	0.87(0.763,0.981)	0.024	-0.22	0.81(0.711,0.914)	0.001

Coef =coefficient, HR=*Hazard ratio, 95%CI= 95% confidence interval for Hazard ratio,*

Ref = Reference

Table 4. 4: Multi-variable Analysis of Age at First Marriage and First Sexual Intercourse using Cox PH Model

Variables	Age at First Marriage			Age at First Sexual Intercourse		
	Coef	HR (95%CI)	p-value	Coef	HR (95%CI)	p-value
Region						
Tigray	Ref					
Affar	-0.16	0.86(0.762,0.962)	0.009	-0.36	0.7(0.619,0.781)	0.000
Amhara	0.43	1.54(1.421,1.677)	0.000	0.18	1.2(1.106,1.305)	0.000
Oromiya	-0.14	0.87(0.788,0.963)	0.007	-0.33	0.72(0.648,0.791)	0.000
Somali	-0.24	0.79(0.691,0.897)	0.000	-0.46	0.63(0.557,0.723)	0.000
Benishangul-G	0.18	1.19(1.072,1.328)	0.001	-0.04	0.96(0.868,1.073)	0.508
SNNP	-0.36	0.69(0.624,0.773)	0.000	-0.55	0.58(0.518,0.641)	0.000
Gambela	0.19	1.21(1.071,1.357)	0.002	0.13	1.14(1.011,1.277)	0.032
Harari	-0.02	0.98(0.851,1.139)	0.833	-0.25	0.78(0.677,0.905)	0.001
Dire Dawa	-0.45	0.64(0.550,0.743)	0.000	-0.67	0.51(0.440,0594)	0.000
Women education Level						
No education	Ref					
Primary	-0.69	0.5(0.477,0.531)	0.000	-0.69	0.5(0.476,0.529)	0.000
Secondary	-1.72	0.18(0.143,0.225)	0.000	-1.62	0.2(0.161,0.246)	0.000
Higher	-1.63	0.2(0.145,0.266)	0.000	-1.26	0.28(0.217,0.373)	0.000
Mass media						
No	Ref					

Yes	0.04	1.04(0.999,1.126)	0.11			
Wealth index						
Poor	Ref					
Middle	0.01	1.01(0.991,1.094)	0.637	0.02	1.02(0.964,1.062)	0.532
Rich	0.06	1.06(0.998,1.124)	0.055	0.08	1.08(1.025,1.146)	0.005
Women work Status						
Not had work	Ref					
Had work	0.006	1.07(0.959,1.056)	0.78	0.01	1.01(0.964,1.062)	0.645
Religion						
Orthodox	Ref					
Muslim	0.02	1.02(0.943,1.096)	0.665	0.000	1(0.927,1.078)	0.994
Protestant	-0.03	0.97(0.886,1.053)	0.428	-0.09	0.92(0.839,0.997)	0.044
Others	-0.02	0.98(0.854,1.129)	0.802	-0.09	0.91(0.793,1.047)	0.19

Coef =coefficient, HR=*Hazard ratio, 95%CI= 95% confidence interval for Hazard ratio,*

Ref = Reference

In the Multivariable analysis of Cox PH model (table 4.4 above) only region and women education level were significant at 5% level of significance while Mass media, wealth index, women work status and religion were not significant at 5% level of significance for time to first marriage. Region and women's level of education when no education is as reference statistically significantly prolong time-to-first marriage among women in Rural Ethiopian.

With regard to time to sexual intercourse region, women education level, wealth index and religion were significant at 5% level of significance. Women work status was not significant at 5% level of significance. Region and women's level of education when no education is as reference statistically significantly prolong time-to-first sexual intercourse while wealth index (poor is reference) shorten the survival time for time-to-age at first marriage among women in Rural Ethiopian. From the wealth index category middle were not significant when poor is reference (p=0.532).

4.2.1 Model Diagnostics

The Cox proportionality assumption was tested by including an interaction term of log-transform of time with each covariate to be tested. Table 4.5 presents the results of model checking for the Cox PH model. The result showed that time-dependent covariates (interaction of covariates with

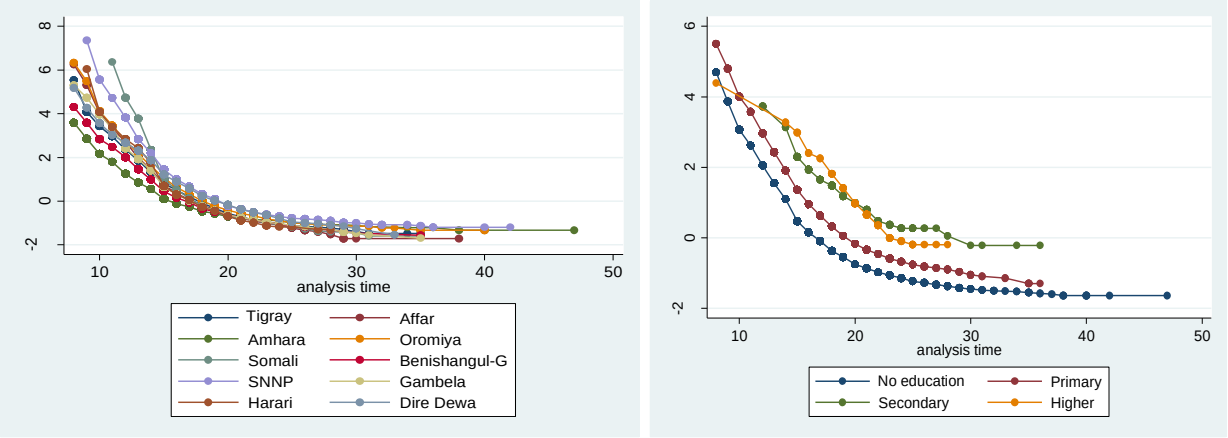
logarithm of time) were significant which justifies the proportional hazard assumption violates at 5% level of significance.

Table 4. 5: Cox Proportional Odds Assumption test for the variables included in the multivariable model

Variables	Age at First Marriage			Age at First Sexual Intercourse		
	Rho	Chisq	p-value	Rho	Chisq	p-value
Region	0.022	4.12	0.042	0.019	4.233	0.021
Education	-0.124	206.33	0.00	-0.114	168.668	0.000
Wealth				0.034	9.782	0.001
Religion				0.003	4.101	0. 047
Region*log(time)	-0.023	4.32	0.037	-0.019	5.184	0.021
Education*log(time)	0.126	211.12	0.000	0.115	171.763	0.000
Wealth*log(time)				-0.035	9.99	0.001
Religion*log(time)				-0.004	5.121	0.047
GLOBAL	NA	223.38	0.000	NA	183.352	0.000

We also used the – (log (-log (survival))) versus survival time plot to check the PH assumption for all the categorical variables included in the model. The result is presented in figure 4.3 below The graphs for each of the categorical variable display lines that are not be parallel implying that the proportional-hazards assumption among categorical variable such as region, women educational level, wealth index and religion has been violated. When the proportional hazards assumption is not satisfied, the Cox proportional hazards model would not be suitable. In this case, we used an accelerated failure time models (AFT), which is an alternative model when proportional hazard assumptions are not satisfied.

Plot of –log(-log(survival)) versus time to assess the proportionality assumption for region and Women education level for age at first marriage



Plot of –log(-log(survival)) versus time to assess the proportionality assumption for religion and Women education level for age at first sexual intercourse

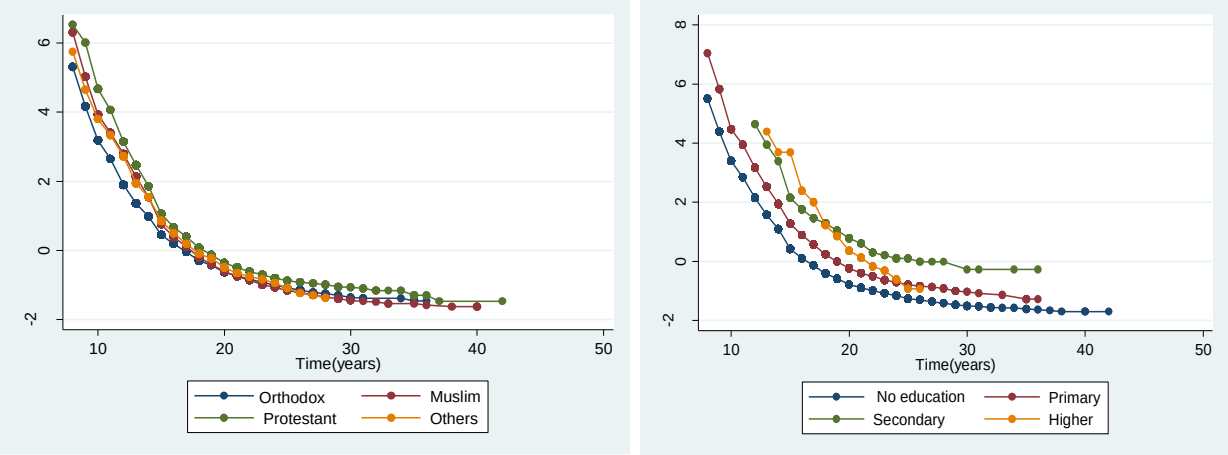


Figure 4. 3: Test for Proportional Odds Assumption for the Cox Model

Cox- Snell Residuals Plots

The Cox-Snell residual is used to check the overall fit of the data. The plot for the fitted model of residuals for the Cox PH model to our data is given in figure 4.4 below. If the model fits the data, the plot of cumulative hazard function of residuals against Cox-Snell residuals should be approximately a straight line with slope 1. We can see that the plot of the cumulative hazard

function against Cox-Snell residuals is nearest to the 45° straight lines through the origin. This suggests that Cox PH model is fitted age at first marriage and first sexual intercourse datasets.

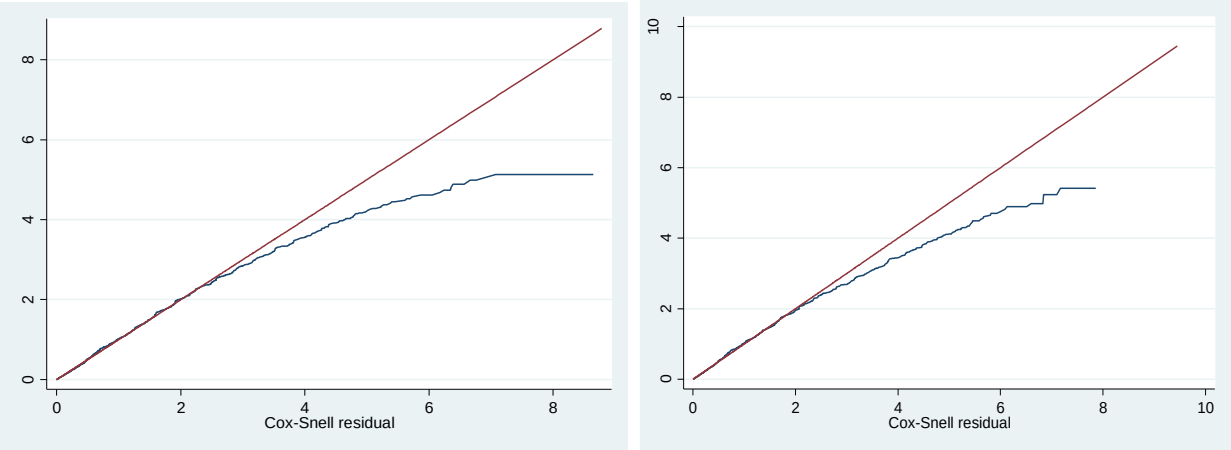


Figure 4. 4: Cox- Snell residuals plots for AFM and AFSI data respectively

4.3 Accelerated Failure Time Model

The AFT model which is another alternative of the Cox PH model when the PH assumption is violated can be used to express the magnitude of effect in a more accessible way in terms of difference between covariates in survival time. We fitted the datasets using Weibull, log-logistic and log-normal AFT model.

4.3.1 Univariable Analysis

The univariable analyses was fitted for every covariate by AFT models using different baseline distributions i.e. weibull, log-logistic, and log-normal. In all univariable analysis of AFT models, region, wealth index of the family, women educational level, women's work status and religion were significant at 20% level for all the models considered, while mass media was not significant for the two responses and women work status was not significant for time to first sexual intercourse of log-normal model. The summary of univariable analysis is given in Appendix A. Therefore those variables that are significant in univariable analysis were included in the final model (multivariable analysis).

4.3.2 Multivariable AFT Model

For age at first marriage and firs sexual intercourse data, multivariable AFT models of weibull, log-logistic, and log-normal distributions were fitted by including all the covariates those are significant in the univariable analysis at 20% level. To compare the efficiency of different models, the AIC was used. A model having the minimum AIC value was preferred. Accordingly, Log-logistic AFT model (AIC = 46,830.22) found to be the best for age at first marriage and age at first sexual intercourse (AIC = 46,031.71) data sets from the given alternatives when we include all the covariate those are significant in the univariable analysis. AFT models and the corresponding AIC values are displayed in Table 4.6.

Table 4. 6: Comparison of AFT Models Using AIC criteria for AFM and AFSI data

Baseline Distribution	Age at First Marriage	Age at First Sexual Intercourse
	AIC	AIC
Weibull	49,916.13	49,235.79
Log- logistic	46,830.22	46,031.71
Log- normal	47,228.63	46,405.15

AIC=Akaike's information criteria

The output of the log-logistic AFT model is presented in Table 4.7. Women's educational level when using no education as reference and religion (Orthodox as reference) prolong the survival time for time-to-age at first marriage and first sexual intercourse, while wealth index when using poor as reference and some categories of region when Tigray was reference statistically significantly shorten time-to-age at first marriage and first sexual intercourse in Rural Ethiopia.

Under the log-logistic AFT model, from region category Affar region (p-value =0.181, ϕ=1.02, 95% CI: (0.993,1.039)) and Harari region(p-value=0.949, ϕ=1.001,95% CI:(0.9731,1.0288)) are not significant when compare to Tigray region for time to first marriage but for time to first sexual intercourse Benishangul-Gumuz region(p-value=0.524, ϕ=0.99,95% CI:(0.975,1.013)) was not significant when compare to Tigray region. From the variable religion category Muslim was not significant when using orthodox as the reference category for time to marriage and sexual intercourse with (p-value=0.081, ϕ = 1.01, 95%CI: (0.998, 1.028)) and (p-value=0.148, ϕ = 1.01, 95%CI: (0.997, 1.023)) respectively. Also, households with middle wealth index was

insignificant for first marriage and first sexual intercourse by using poor households as a reference category with (p-value =0.292, 95%CI: (0.983,1.005), ϕ =0.99) and (p-value =0.552, 95%CI: (0.986,1.008)and ϕ =0.99) respectively.

Table 4. 7: Log-logistic Multi-variable AFT Model for AFM and AFSI data

Variables	Age at First Marriage			Age at First Sexual Intercourse		
	Coef	ϕ (95%CI)	p-value	Coef	ϕ (95%CI)	p-value
Region						
Tigray	Ref					
Afar	0.01	1.02(0.993,1.039)	0.181	0.06	1.06(1.041,1.086)	0.000
Amhara	-0.1	0.89(0.877,0.906)	0.000	-0.05	0.95(0.937,0.966)	0.000
Oromiya	0.02	1.02(1.001,1.041)	0.036	0.06	1.06(1.046,1.084)	0.000
Somali	0.05	1.05(1.022,1.075)	0.000	0.09	1.09(1.070,1.121)	0.000
Benishangul-G	-0.06	0.95(0.926,0.965)	0.000	-0.006	0.99(0.975,1.013)	0.524
SNNP	0.07	1.08(1.054,1.099)	0.000	0.11	1.12(1.096,1.139)	0.000
Gambela	-0.05	0.95(0.925,0.969)	0.000	-0.03	0.96(0.946,0.987)	0.002
Harari	0.001	1.001(0.974,1.029)	0.949	0.05	1.05(1.023,1.077)	0.000
Dire Dawa	0.09	1.09(1.060,1.126)	0.000	0.13	1.14(1.113,1.177)	0.000
Women education level						
No education	Ref					
Primary	0.1283	1.14(1.126,1.148)	0.000	0.12	1.13(1.119,1.140)	0.000
Secondary	0.3526	1.42(1.369,1.478)	0.000	0.32	1.38(1.335,1.429)	0.000
Higher	0.3762	1.46(1.380,1.540)	0.000	0.29	1.34(1.284,1.407)	0.000
Wealth index						
Poor	Ref					
Middle	-0.0062	0.99(0.983,1.005)	0.292	-0.003	0.99(0.986,1.008)	0.552
Rich	-0.0246	0.98(0.965,0.987)	0.000	-0.023	0.98(0.967,0.987)	0.000
Women work status						
Not had work	Ref					

Had work	-0.004	0.99(0.986,1.006)	0.419			
Religion						
Orthodox	Ref					
Muslim	0.01	1.01(0.998,1.028)	0.081	0.0098	1.01(0.997,1.023)	0.148
Protestant	0.02	1.02(1.006,1.039)	0.009	0.0265	1.03(1.011,1.043)	0.001
Others	0.01	1.01(0.986,1.043)	0.328	0.0209	1.02(0.995,1.048)	0.119

Coef =coefficient, ϕ=acceleration factor, 95%CI= 95% confidence interval for acceleration factor, Ref = Reference

From table 4.7: the estimated acceleration factor for women from Amhara, Oromiya, Somali, Benishangul-Gumuz, SNNP, Gambela and Dire Dawa is estimated to be 0.89, 1.02, 1.05, 0.95, 1.08, 0.95, 1.09 with 95% CI; (0.877,0.906), (1.001,1.041), (1.022,1.075), (0.926,0.965), (1.054,1.099), (0.925,0.969), (1.06,1.126) respectively by using Tigray region as reference category. This indicates women from Oromiya**,** Somali**,** SNNP and Dire Dawa have prolonged survival of time to age at first marriage and the survival of time-to-age at first marriage decreased in Amhara, Benishangul-Gumuz and Gambela region than Tigray region women.

In the case of time to first sexual intercourse the acceleration factor for women from Affar, Amhara, Oromiya, Somali, SNNP, Gambela, Harari and Dire Dawa is estimated to be 1.06, 0.95, 1.06, 1.09, 1.12, 0.96, 1.05 ,1.14 with 95% CI; (1.041,1.086), (0.937,0.966), (1.046,1.084), (1.07,1.121), (1.096,1.139), (0.946,0.987), (1.023,1.077), (1.113,1.177) respectively by using Tigray region as reference category. This indicates women from Affar, Oromiya**,** Somali**,** SNNP, Harari, and Dire Dawa have prolonged survival of time to age at first sexual intercourse and the survival of time-to-age at first sexual intercourse decreased in Amhara and Gambela than Tigray women.

The acceleration factors for women attending primary education, secondary education, and higher are 1.14, 1.42, and 1.46 with 95% CI; (1.126, 1.148), (1.369, 1.478), (1.38, 1.54) respectively. These confidence intervals does not include one in all category ; indicating primary, secondary and higher education's were significantly important factors for the timing of age at first marriage by using uneducated women as a reference category. This indicates that women

with primary, secondary and higher education prolonged the age at first marriage, and also education prolongs the survival of time to first sexual intercourse by a factor of 1.13, 1.38, and 1.34 with 95% CI: (1.119,1.140), (1.335,1.429), (1.284,1.407) for primary, secondary and higher education respectively.

The acceleration factors for middle wealth index and rich are 0.99 and 0.98 with 95% CI; (0.983, 1.005), and (0.965, 0.987) respectively by using poor household as reference. This implied that poor house hold women have longer survival of time-to-age at first marriage, however the difference is not significant for middle and poor house hold wealth index (P=0.292).Similarly poor household women has longer survival time than the rich household for first sexual intercourse. The acceleration factor for protestant religion was1.02 with 95% CI: (1.006, 1.04) by using Orthodox as reference category. This result suggested that women from Protestant religion had longer survival of time-to-age at first marriage than women from Orthodox religion. Women form protestant religion also prolong survival time to first sexual intercourse by a factor of 1.03 with 95% CI :(1.011, 1.043) compared to orthodox religion.

4.3.3 Model Diagnostics: AFT Model

Checking Adequacy of Parametric Baselines

After the model has been fitted, it is desirable to determine whether a fitted parametric model adequately describes the data or not. Therefore, the appropriateness of model with weibull baseline can be graphically evaluated by plotting log (-log(S(t)) versus log(time), the log-logistic baseline by plotting $\log(\frac{1-\hat{S}(t)}{\hat{S}(t)})$ versus log(time) and the log-normal baseline by plotting $\Phi^{-1}[1-\hat{S}(t))]$ versus log (t). If the plot is linear, the given baseline distribution is appropriate for the given dataset. Accordingly, their respective plots are given in figure 4.5 below for age at first marriage and age at first sexual intercourse and the plot for the log-logistic baseline distribution make straight line better than weibull and log-normal baseline distribution for the two data sets. This evidence also strengthens the decision made by AIC value that log-logistic baseline distribution is appropriate for the given datasets.

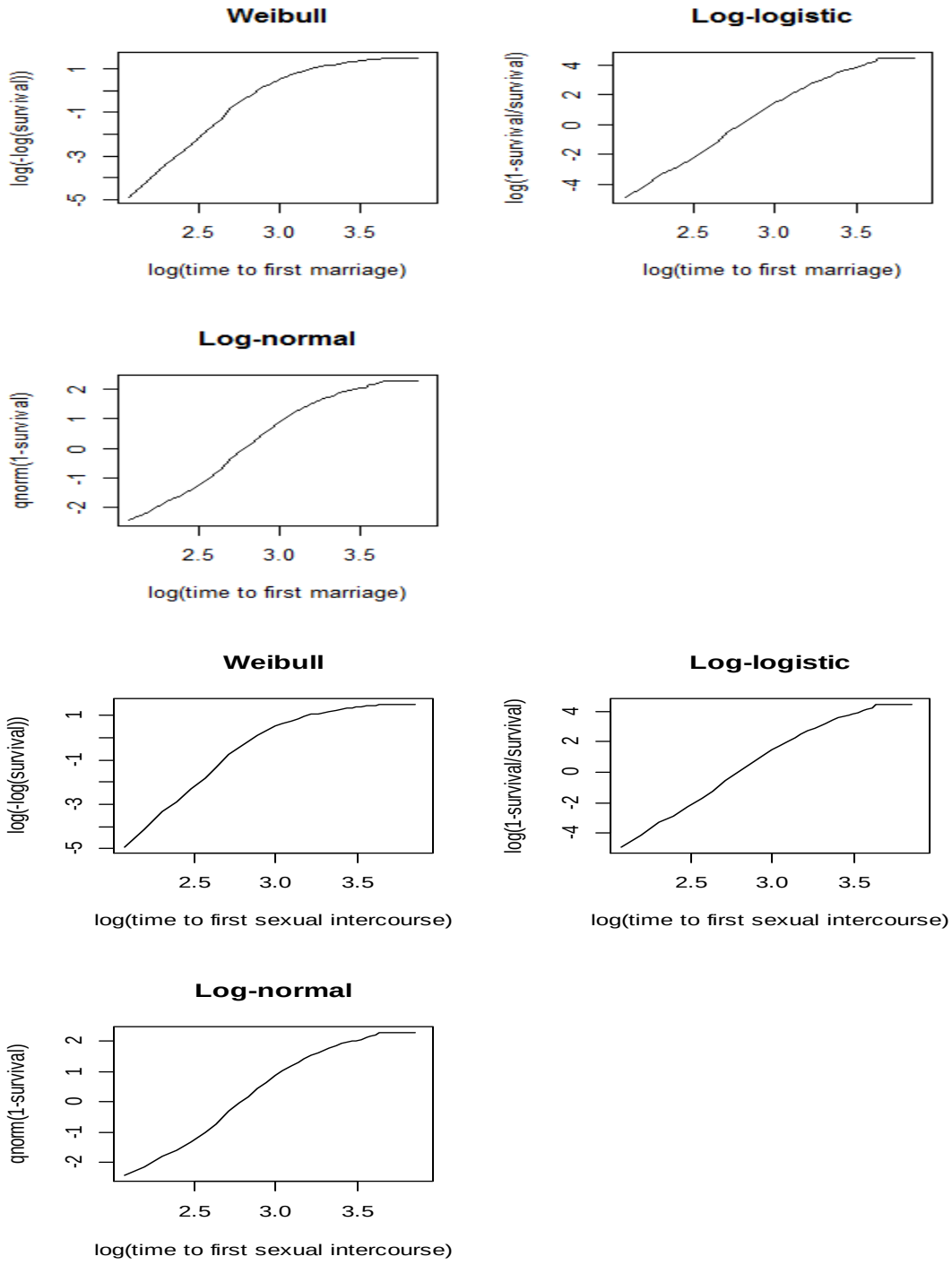


Figure 4. 5: Graphs of Weibull, Log-logistic, and Log-normal baseline distributions for age at first marriage and age at first sexual intercourse data sets

Cox- Snell Residuals Plots

The Cox-Snell residuals are one way to investigate how well the model fits the data. The plot for fitted model of residuals for log-logistic to our data via maximum likelihood estimation with cumulative hazard functions is given in figure 4.6 below. If the model fits the data, the plot of cumulative hazard function of residuals against Cox-Snell residuals should be approximately a straight line with slope 1. We can see that the plot of the cumulative hazard function against Cox-Snell residuals in figure 4.6 is nearest to the 45° straight lines through the origin for log-logistic model suggesting that it is appropriate for age at first marriage and age at first sexual intercourse data sets.

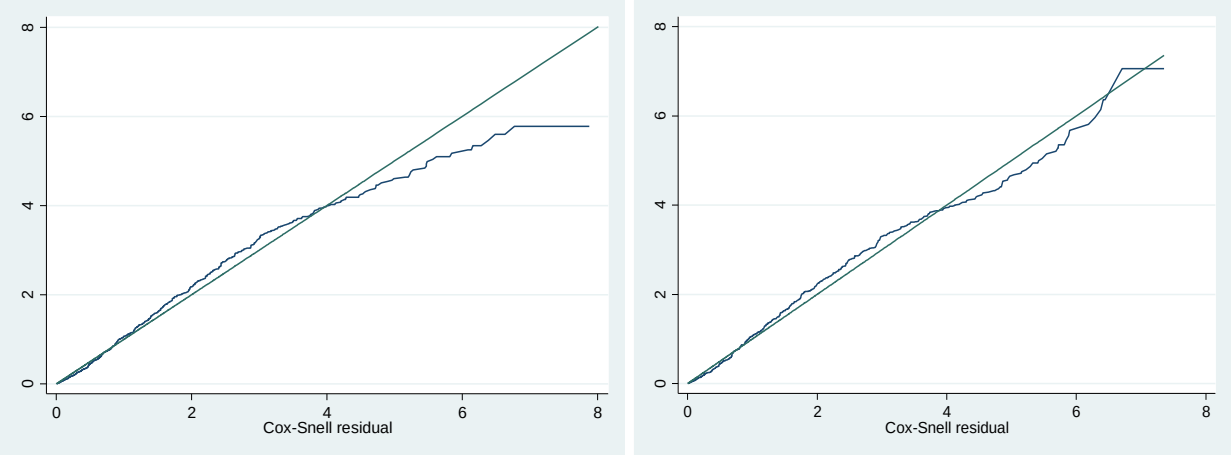
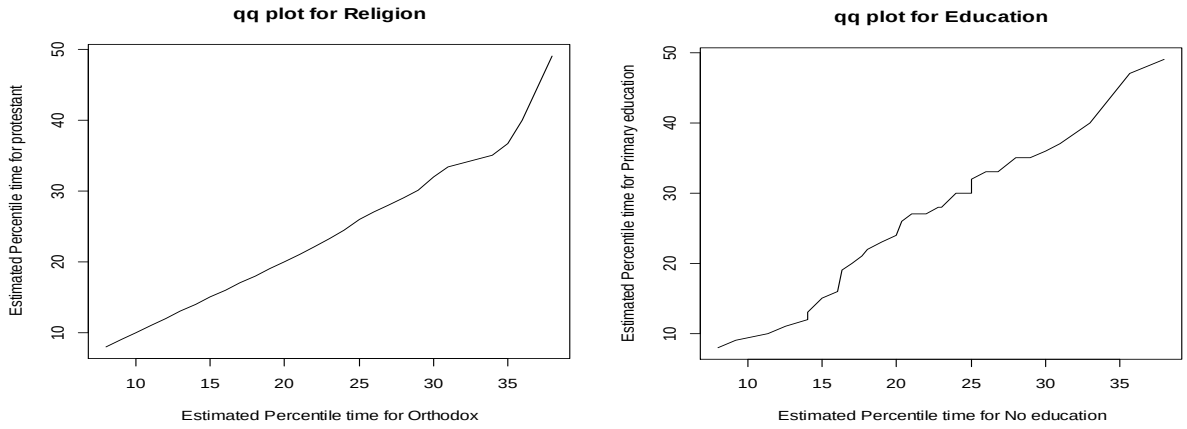


Figure 4. 6: Cox- Snell residuals plots of log-logistic baseline distribution for age at first marriage and age at first sexual intercourse data respectively

Quantile-Quantile Plot

A quantile-quantile or q-q plot is used to check if the AFT provided an adequate fit to the data using by two different groups of population. We shall graphically check the adequacy of the model by comparing the significantly different groups of women education level(no education and primary), as well as religion(Orthodox and protestant) for age at first marriage and sexual intercourse. The figures appear to be approximately linear for both women's educational level and religion as shown in figure 4.7. Therefore the accelerated failure time appears to be the best to describe age at first marriage and age at first sexual intercourse data sets.

Quantile-Quantile Plot for Age at First Marriage data



Quantile-Quantile Plot for Age at First Sexual Intercourse data

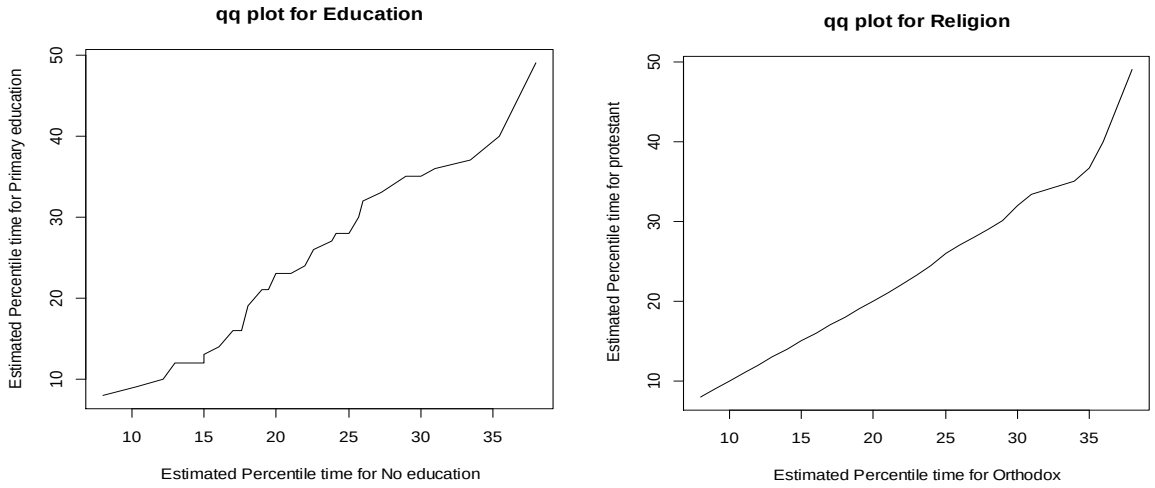


Figure 4. 7: Quantile- Quantile plot to check the adequacy of the accelerated failure time model

4.4 Comparison of Cox PH and Log-logistic AFT Model

In this study, in order to compare the efficiency of the models the AIC was used. From the Table 4.8, we can see that the log-logistic AFT model has a minimum AIC for age at first marriage and age at first sexual intercourse than Cox PH model indicating that log- logistic AFT model fitted the survival of time-to-first marriage and first sexual intercourse data better than the Cox PH model which did not satisfy the proportional hazard assumption.

Table 4. 8: Comparison of Cox PH and log-logistic AFT Models

Model	Age at First Marriage	Age at First Sexual Intercourse
	AIC	**AIC**
Cox PH	136739.1	137,968.50
Log-logistic	46,830.22	46,031.71

4.5 Discussion

The main goal of the study was modeling the determinants of both age at first marriage and age at first sexual intercourse among women in Rural Ethiopia separately using Cox PH and AFT models by considering three baseline distributions: Weibull, log-logistic, and log-normal distributions. Covariate which were included in the study were region, women educational level, mass media, wealth index, women work status and religion and the outcome variable of interest were the survival of age first marriage and age at first sexual intercourse measured in years.

The univariable analysis (in Appendix A) revealed that region, women educational level, wealth index of the family, women work status and religion were significantly associated with age at first marriage but access to mass media was not significant at 20% level. All significant covariates in univariable analysis were included in all multivariable analysis of AFT model and comparison was done within the models using AIC criteria where the model having minimum AIC value is selected to be the best (Munda, 2012). Log-logistic AFT model was found to be the best over Weibull and Log-normal AFT models based on AIC value and graphical evidence (figure 4.5).Women work status had no significant association to the survival of time-to-age at first marriage while the covariates region, women's educational level, and wealth index were significantly associated with the timing of age at first marriage. Finally the two models, Cox PH and log-logistic AFT, were compared based on AIC, suggesting that log-logistic AFT model fits the data better than Cox PH model which did not satisfy the proportional hazard assumption.

The findings of this study revealed that wealth index and from the category of region Amhara, Benishangul-Gumuz and Gambela region were significantly shorten the time-to-age at first marriage while women's educational level, religion, and from region category(Oromiya,SNNP

and Dire Dawa) prolonged time-to-age at first marriage among women in Rural Ethiopia. About 67.1% of women were married before age of 18 years. This indicates that early marriage is highest in Ethiopia The median survival time of age at first marriage for Rural Ethiopian women was 17 years with 95% CI; (16.90, 17.10).

The findings of this study suggested that women's educational level had a significant effect on the survival time of age at first marriage with 5% level of significance and it prolonged age at first marriage by the factor of $\phi = 1.14$, $\phi = 1.42$ and $\phi = 1.46$ for primary, secondary and higher education respectively when compared to illiterate women. The result of the study shows that woman who had higher education was more survived than those uneducated and primary education. A similar study conducted in Ethiopian by Tezera (2013) used data from 2011 EDHS to examine the effect demographic and socioeconomic variables to determine early marriage among women. The result of the study showed that educational level of women was found to be significant predictors for early marriage. This finding is also supported by Manda and Meyer (2005) who found that women with higher levels of educational attainment are far more likely to enter marriage at later age than those without any or with little education. Similar study in Malawi by Palamuleni (2011) also found that women education were statistically significant factor of age at first marriage. Also, Agaba et al (2011) indicates that educational attainment were strong socio-economic determinants of first marriage in Western Uganda. A similar study in Nigeria by Adebowale *et al*. (2012) showed that education were significantly associated with timing of age at first marriage. This study also indicated that age at first marriage of women was directly related to levels of education i.e., Woman who had primary, secondary and higher education was less likely to marry earlier than those with no education. Study on Socio-economic determinants of age at first marriage of the Ethnic Tribal women in Bangladesh by Kamal (2011) also revealed that woman's educational attainment was significantly delayed the timing of marriage. Another study in Kenya by Ikamari LD, (2005) also investigated the effect of education on the timing of marriage. The result showed that education has a statistically significant and strong positive effect on a woman's age at first marriage; the effect remaining robust in the presence of a number of controls. The highly educated women are more likely to delay marriage. Significant variations in the effect of education across the generations of women

are apparent. The effect is greater for the younger women, indicating increased postponement of marriage.

The results of this study suggested that region was significant predictive factor for the timing of age at first marriage of women in Rural Ethiopia. Women in the Oromiya region, Somali region, SNNP region and Dire Dawa administration prolong age at first marriage by a factor of ϕ =1.02, ϕ =1.05, ϕ =1.08, and ϕ =1.09 respectively compared to those in the Tigray region. However, women from Amhara region (ϕ =0.89), Benishangul-Gumuz region (ϕ =0.95) and Gambela region (ϕ =0.95) had a significantly higher risk of early first marriage compared to their counterparts in the Tigray region. A study conducted by Tezera (2013) to examine the effect demographic and socioeconomic variables to determine early marriage among women in Ethiopia. The result of the study revealed that region of women was significant predictor for early marriage. A similar study in Nigeria by Adebowale *et al.* (2012) also found that region was significantly associated with the timing of first marriage. Also another study in Vietnam by Lung Vu (2009) finds that region was significantly related to age at first marriage. This finding is consistent with Ikamari LD, (2005) in Kenya, Palamuleni (2011) in Malawi, Erulkar (2013) in Ethiopia.

The result of this study revealed that household wealth index were significantly shortened the survival of time to age at first marriage in Rural Ethiopian women. Compared to rich women, poor women had relatively lower risk of first marriage. This result contradicts with a study by Kamal (2011) in the Ethnic Tribal women in Bangladesh. The findings revealed that Parents' economic status had the most significant effect on marital timing. The higher the economic status of the parents, the lower is the probability of age at earlier marriage. Studies elsewhere suggest the opposite of this, with poor women having a relatively higher risk of first marriage (Haloi and Limbu, 2013; Hoq,2013).

The result of this study also revealed that religion was an important factor for age at first marriage of women in Rural Ethiopia. Women from protestant religion had prolonged age at first marriage by a factor of ϕ =1.03 as compared to Orthodox. These findings are similar to Tezera (2013) in Ethiopia and the study showed that religion of women was found to be significant predictors for early marriage. The same study in Nigeria by Adebowale *et al*. (2012) found

religion were significantly associated with the timing of first marriage. Another study by Agabaet al (2011) religion is strong social determinant of first marriage in Western Uganda. This result is consistent with Erulkar (2013) in Ethiopia, Abdallah (2011) in Nigeria, Bayisenga (2012) in sub-Saharan Africa, Ikamari LD, (2005) in Kenya.

With regard to age at first sexual intercourse, the univariable analysis (Appendix A) revealed that region, women's educational level, wealth index of the family, and religion were significantly associated with age at first sexual intercourse but access to mass media and women work status were not significant at 5% level of significance. All significant covariates in univariate analysis were included in all multivariable analysis of AFT model. Within the models using AIC Log-logistic AFT model was found to be the best over Weibull and Log-normal AFT models based on AIC value and graphical evidence (figure 4.5). The covariates region, women educational level, wealth index, and religion were significantly associated with timing of age at first sexual intercourse.

The findings of this study revealed that wealth index and from the category of region Amhara and Gambela region were significantly shorten the time-to-age at first sexual intercourse while women educational level, religion, and Affar, Somali, Oromiya, SNNP, Harri and Dire Dawa from region category prolonged time-to-age at first sexual intercourse among women in Rural Ethiopia. About 68.6% of the women have had first sexual intercourse before age of 18 years. This indicates that early first sexual intercourse is highest in Ethiopia. 61.4% of the women have had first sexual intercourse at their first union. This suggests that Rural Ethiopian women begin sexual intercourse at the time of their first marriage. The median survival time of age at first sexual intercourse for Rural Ethiopian women was 17 years with 95% CI; (16.90, 17.11).

The findings of this study suggested that the women educational level had a significant effect on the survival time of age at first sexual intercourse at 5% level of significance and it prolonged age at first marriage by the factor of $\phi = 1.13$, $\phi = 1.38$ and $\phi = 1.34$ for primary, secondary and higher education respectively when compared to illiterate women. The result of the study shows that woman who had secondary education was more survived than those uneducated and primary education. Recent analyses of DHS data have shown that increased educational attainment in sub-Saharan Africa has contributed to the decline in early marriage (Mensch et al. 2005) and that

girls who are enrolled in school are less likely than those who are not enrolled to engage in premarital sex (Blanc et al. 2005).

The results of this study suggested that region was significant predictive factor for the timing of age at first sexual intercourse of women in Rural Ethiopia. Women in the Affar region, Oromiya region, Somali region, SNNP region, Harari region and Dire Dawa administration prolong time to age at first sexual intercourse by a factor of ϕ =1.06, ϕ =1.06, ϕ =1.09, ϕ =1.12, ϕ=1.05, and ϕ=1.14 respectively compared to those in the Tigray region. However, women from the Amhara region (ϕ =0.95) and Gambela region (ϕ =0.96) had significantly higher risk of early first sexual intercourse compared to their counterparts in the Tigray region.

The result of this study revealed that household wealth index were significantly shortened the survival of time to age at first sexual intercourse in Rural Ethiopian women. Compared to rich women, poor women had relatively lower risk of first sexual intercourse.

The result of this study also revealed that religion was an important factor for age at first sexual intercourse of women in Rural Ethipia. Women from protestant religion had prolonged age at first marriage by a factor of ϕ =1.02 as compared to Orthodox. A recent study in Kenya by Okigbo CC(2015) found that Religion were associated with time to first sex.

5 Conclusions and Recommendations

5.1 Conclusions

This study was used age at first marriage and age at first sexual intercourse datasets among women in Rural Ethiopia which were obtained from central statistics agency with an aim of modeling the determinant of age at first marriage and age at first sexual intercourse by using Cox PH and AFT models. Out of the total 10,417 women, about79.4% were married while 20.6% of them were not married and with regard to sexual intercourse, 80.1% of women had sexual intercourse, 19.9% were not had sexual intercourse at the time of the survey. The median time for both age at first marriage and age at first sexual intercourse were 17 years. This indicates that Rural Ethiopian women had first sexual intercourse at the time of their first union.

To model the determinants of age at first marriage and age at first sexual intercourse, Cox PH and different AFT models by using different baseline distributions were applied. Among this using AIC, log-logistic AFT model is better fitted to age at first marriage and first sexual intercourse datasets than Cox PH and other AFT models.

The result of Log-logistic AFT model showed that region, women's educational level, wealth index and religion were found significant predictors to age at first marriage and first sexual intercourse among women in Rural Ethiopia. Among these significant predictors, women's educational level and religion of women prolong age at first marriage and first sexual intercourse while wealth index of the family shortens timing of first marriage and first sexual intercourse.

Goodness of the fit of baseline distribution by means of graphical method and Cox-Snell residuals plots in figure 4.5 and 4.6 revealed that log-logistic distribution is better when compared to Weibull and log-normal baseline distributions to explain time-to-first marriage and first sexual intercourse datasets.

5.2 Recommendations

Based on the results obtained in the analysis of this thesis work, the following recommendations are made for policy makers and the community at large.

I. Awareness about the importance of elongating both age at first marriage and age at first sexual intercourse should be given for rural women by the government and other concerned bodies.

II. The study findings reveal that education is the most significant variable affecting age at first marriage and age at first sexual intercourse in rural Ethiopia. Therefore, it is important that government policies promote the status of women in rural Ethiopia by helping them to have more access to education so that they can make their own decision regarding when to get married and had sexual intercourse. It is crucial to continue improving girls and young women access to education is important for rising the women's age at first marriage and age at first sexual intercourse, which is vital for empowering them and enhancing their participation in any sector. The education system should aim at providing life skills to enable girls avoid early marriage and early sexual debut as well as providing reproductive health information so that they are aware of the advantages of delayed marriage and sexual intercourse.

III. It is important that the government ensure that the provision of services like education, health and other opportunities are done throughout the country with more emphasis placed on the parts of the country where early marriage and sexual intercourse is highest.

IV. Further studies should be conducted in each region of Ethiopia and identify other factors that are not identified in this study. Based on that study, regional governments should take actions to elongate time-to-first marriage and first sexual intercourse.

References

Aalen, O., Borgan, O., & Gjessing, H. (2008), Survival and event history analysis: a process point of view. Springer Science & Business Media.

Abdallah, B. A. (201l): Girl Child Marriage and Women Development in Nigeria: Contemporary Issues", Journal of Development and Psychology 14(9): pp.248-259.

Adebowel A., Fagbamigbe A., Okareh O. and Lawal O. (2012). Survival Analysis of Timing of First Marriage among Women of Reproductive age in Nigeria: *African Journal of Reproductive Health.*

Adedokun, G. N., Tochukwu, H. E. and Adedeji, 0. 0. (2012): "Early Childhood Marriage and Early Pregnancy as a Risk to Safe Motherhood", A Report on the Regional Conference on Traditional Practices Affecting the Health of Women and Children in Africa, 19-20 Nov. ICA.

Agaba. P, Atuhaire. L.K, Rutaremwa. G (2011), Determinants of Age at First Marriage among women in Western Uganda. SSAE, Makerere University, Uganda.

Ango, R. G.(1991): "The. Impact of Girls' Education on Early Marriage", Independent Evaluation Report: CAMIFED and CAMA Programmes Zimbabwe, September.

Assefa, D., Wassie. E., Getahun, M., Berhaneselassie, M., Melaku, A. (2005). *Harmful traditional practices: For the Ethiopian Health Center Team*. Ethiopia Public Health Training Initiative, Awassa College.

Bayisenge, J. (2012):'"Early Marriage as a Barrier to Girl's Education: A Developmental Challenge in Africa", Journal of Social Psychology 12(6): pp. 23-48.

Blanc, A. K., & Way, A. A. (1998).Sexual behavior and contraceptive knowledge and use among adolescents in developing countries. Studies in Family Planning, 29: 106–111.

Blossfeld Hans-Peter,Jaeniches Ursula.(1992). Educational Expansion and Changes in Women's Entry into Marriage and Motherhood in the Federal Republic of Germany, Journal of Marriage and the Family 54 (2):302-315.

Bongaarts, j., cleland, j., townsend, j., bertrand, j. & gupta, m. 2012. Family Planning Programs For the 21ST Century. New York: Population Council.

Central Statistical Authority and ORC Macro. (2001).*Ethiopia demographic and health survey 2000.* Addis Ababa,Ethiopia and Calverton, Maryland, USA: Central StatisticalAuthority and ORC Macro .

Central Statistical Agency and ORC Macro .(2006). *Ethiopia demographic and health survey 2005.* Addis Ababa, Ethiopia and Calverton, Maryland, USA:Central Statistical Agency and ORC Macro.

Central Statistical Agency, (2011). Ethiopian Demographic and Health Survey, Addis Ababa,Ethiopia.

Clark, S. (2004). "Early Marriage and HIV Risks in Sub-Saharan Africa," *Studies in FamilyPlanning,35(3),*149–160.[Online].Available: http://www.ncbi.nlm.nih.gov/pubmed/15511059.

Collett, D. (2003). Modeling survival data in medical research. CHAPMAN & HALL/CRC.

Cox, D. R.(1972). Regression Models and Life Tables (with discussion), Journal of the Royal Statistical Society, Series B, 34(2).

Cox, D.R., Oakes, D. 1984. Analysis of Survival Data. London: Chapman and Hall.

Dress, K.A. (1986). The effect of gender identify on conversation. SocialPsychology Quarterly, 49, 294-301.

Duchateau, L. and Janssen, P. (2008). The Frailty Model. Springer, New York, 1 edition.

Erulkar, A. (2013). Early marriage, marital relations and intimate partner violence in Ethiopia. *International Perspectives on Sexual and Reproductive Health*, 39(1), pp.6-13.

Erulkar, A.S., T. Mekbib, N. Simie, and T. Gulema. 2004. "The experience of adolescence in rural Amhara region, Ethiopia." Accra: Population Council.

Febriany, Vita. 2006. "Community Synthesis Report: Semampir". Jakarta: SMERU Research Institute.

GOLDIN, C. & KATZ, L. F. 2000. The power of the pill: oral contraceptives and women's career and marriage decisions. National Bureau of Economic Research.

Ghosh, Biswajit. 2011. "Child Marriage and Its Prevention: Role of Adolescent Girls". Indian Journal of Development Research and Social Action, Vol. 7, No. 1-2, pp. 49-62

Hallett, T.B., Lewis, J.J., Lopman, B.A., Nyamukapa,C.A., Mushati, P., Wambe, M., Garnett, & Gregson,S. (2007).Age at first sex and HIV infection in ruralZimbabwe. Studies in Family Planning, 38:1-10.

Haloi, A. and Limbu, D.K. (2013). Socio-economic factors influencing age at first marriage of Muslim women of a remote population from North East India, *Antrocom Online Journal of Anthropology* 9 (1) 75-79.

Hertrich, Véronique. (2002). "Nuptiality and gender relationships in Africa: An overview of first marriage trends over the past 50 years," session on Family Change in Africa and Latin America, Population Association of America Annual Meeting, Atlanta, 9–11 May.

Hoq, M. (2013).Regional differentials in age at first marriage among women in Bangladesh, *Asian Journal of Applied Science and Engineering* 2 (2)76-83.

Ikamari LD, (2005). The effect of Education on the Timing of Marriage in Kenya. DemographicResearch volume 12, article 1, pp 1-28: Feb. 2005.

Jejeebhoy SJ (1995). Women's Education, Autonomy, and Reproductive Behavior: Experience from Developing Countries. Oxford: Clarendon Press.

Jones GW (2007). "Delayed Marriage and Very Low Fertility in Pacific Asia." Popul. Dev. Rev., 33(3): 453-478.

Jones, Gavin W. 2010. "Changing Marriage Patterns in Asia". Asia Research Institute Working Paper Series No. 131.

Joseph N., M.Fajar R. and Mayang R. (2012). Prevalence of Child Marriage and Its Determinants among Young women in Indonesia, Child poverty and social protection Conference.*Journal of Statistics 14, 19 - 25*.

Kamal S.M.Mostafa (2011). Socio-Economic Determinants of Age at First Marriage of the Ethnic Tribal Women in Bangladesh, Asian Population studies.

Kaplan, E. L. and Meir, P. (1958). Nonparametric estimation from incomplete observations. Journalof the American Statistical Association, 58(282):457-481.

KAUFMANN, G. & MEEKERS, D. 1998. The impact of women's socioeconomic position on marriage patterns in sub-Saharan Africa. *Journal of Comparative Family Studies,* 29**,** 101-114.

Kleinbaum D, Klein M. (2005). Survival analysis: A self- learning text, New York,

Springer-Verlag.
Korosteleva, O. (2009). Clinical Statistics: Introducing Clinical Trials, Survival Analysis, and Lon-gitudinal Data Analysis. Jones and Bartlett, Otario, 1 edition.

Lawless JF. (1998). Parametric models in survival analysis. In encyclopedia of Biostatistics.Wiley: New York; 3254-64.

Legrand, C. (2005). Heterogeneity in multicenter clinical trials using the frailty model. PhD thesis,Universiteit Hasselt, Campus Diepenbek.

Lesthaeghe R, Kaufmann G, Meekers D (1989). The Nuptiality regimes in sub- Saharan Africa. in Lesthaeghe, R. (ed) Reproduction and Social Organization in Sub-Saharan Africa, University Of California Press, Berkeley.

LIVI-BACCI M 2012. Demographic growth: Between choice and constraint. *In:* LIVI-BACCI M (ed.) *A concise history of world populaiton.* 5th Edition ed.

Lung Vu (2009). Age at first marriage in Vietnam: Trends and determinants. Paper presented at Population Association of America (PAA), April 30- May 2.

Manda, Samuel., and Renate Meyer. 2005. "Age at First Marriage in Malawi: A Bayesian Multilevel Analysis Using a Discrete Time-to-Event Model". Journal of the Royal Statistical Society. Series A (Statistics in Society), Vol. 168, No. 2 marital disruptions. American Journal of Sociology, 94, 110-129.

Mathur, S., Greene, M. & Malhotra, A. (2003). *Too Young to Wed: The Lives, Rights and Health of Young Married Girls.* Washington, D.C: International Center for Research on Women (ICRW).

Melah, G.S., A.A. Massa, U.R. Yahaya, M. Bukar, D.D. Kizaya, and A.U. El-Nafaty. 2007. Risk factors for obstetric fistulae in north-eastern Nigeria. Journal of Obstetrics and Gynecology 27(8): 819-23.

Mensch Barbara S., Suchela Singh, and John B. casterline.(2005). Trends in the Timing of First Marriage Among Men and Women in The Developing World.Population council, No.202.

MENSCH, B., GRANT, M. & BLANC, A. 2006. The Changing Context of Sexual Initiation in sub-Saharan Africa. *Popul Dev Rev,* 32**,** 699-727.

Ministry of Finance and Economic Development (MOFED). 2005. Building on Progress: A Plan for Accelerated and Sustained Development to End Poverty (PASDEP), 2005/06–2009/10. N.

Molokwu, B. (2000), “Gender assessment: Issues and Perspectives”, Rwanda ISP Kigali, Rwanda.

Morita, J.G., Lee, T.W. and Mowday, R.T. (1989). Introducing survival analysis toorganizational researchers: A selective application to turnover research. Journal of Applied Psychology, 74, 280-292.

National Committee on Traditional Practices of Ethiopia. (2003). *Ethiopia harmful traditional practices: Old beyond imaginings* Addis Ababa.

National Research Council and Institute of Medicine of the National Academies (NRC-IOM).2005. Growing Up Global: The Changing Transitions to Adulthood in Developing Countries.Ed. Cynthia B. Lloyd. Washington, DC: The National Academies Press. New York.

Palamuleni. E. Martin, (2011). Socioeconomic determinants of age at first marriage in Malawi.International Journal of Sociology and Anthropology vol. 3(7), pp224-235.

Population Council (2004). Supporting Married Girls: Calling Attention to a Neglected Group.

Population Council, (2002). Background document prepared for working group on Girls, United Nations Special Session on Children, 10 may, New York.

Rosenberg, J. 2002. Age at first sex and human papillomavirus infection linked through behavioral factors and partner’s traits. Perspectives on Sexual and Reproductive Health 34(3): 171-172.

Santhya, K., Ram,U., Acharya, R., Jejeebhoy, S., Ram, F., & Singh, A. (2010). *International Perspectives on Sexual and Reproductive Health, 36(3),* 132-139. New York, NY: Population Council.

Sarkar, Prosannajid. 2009. "Determinants and Effect of Early Marriage in Bangladesh, 2007".Research Journal of Applied Sciences 4 (5): 178-184.

Save the Children. (2004). *State of the World’s Mothers 2004.* Westport, CT: Save the Children.

Savitridina, 1997. "Determinants and Consequences of Early Marriage in Java, Indonesia". Asia Pacific Population Journal, (Jun., 1997), 12(2): 25-48.

Shapiro, D. (1996). Fertility Decline in Kinshasa, Population Studies Vol50, No.1, pp 89-

103.

Singh S. and Samara R. (1996), “Early marriage among women in developing countries” in International Family Planning perspectives, vol.22, No 4 pp 148-157+175, Guttmacher Institute.

Johnson-Lans, Shirley., and Patricia Jones. 2011. "Child Brides in Rural India". Vassar College, Working Papers No. 094.

Stevens, V.J. and Hollis, J.F. (1989). Preventing smoking relapse using anindividually tailed skills training technique. Journal of Consulting and ClinicalPsychology, 57, 420-424.

Stover J. 1998. Revising the proximate determinants of fertility framework: what have we learned in the past 20 years? Stud Fam Plann;29:255–67.

Tezera A. (2013). Determinants of Early Marriage among Women in Ethiopia.

UN. (1990).Patterns of First Marriage: Timing and Prevalence, New yorkeVol.4. No. 3 (April), pp: 221-235.

UNFPA. (2006): “In ending child marriage, A guide for global policy action International Planned Parenthood Federation and the Forum on Marriage and the Rights of Women and Girls. U.K.

UNICEF. (2001). *Early Marriage: Child Spouses.* UNICEF Innocenti Research Centre. Florence, Italy: UNICEF.

UNICEF. (2005). *Early Marriage: A Harmful Traditional Practice: A Statistical Exploration.* New York, NY: UNICEF.

UNICEF. (2007)b. *The State of the World's Children 2007.* New York: UNICEF.

UNICEF. (2011). Child Protection from Violence,Exploitation and Abuse.

United Nations, (2000). World Marriage Patterns. New York: Population Division.

WAITE LJ 2006. Marriage and family. *In:* POSTON DL & MICKLIN M (eds.) *Handbook of Population.* New York: Springer.

WELLINGS, K., COLLUMBIEN, M., SLAYMAKER, E., SINGH, S., HODGES, Z., PATEL, D. & BAJOS, N. 2006. Sexual behaviour in context: a global perspective. *The Lancet,* 368**,** 1706-1728.

Welz T., Hosegood V., Hosegood, V., Jaffar, S., Batzing- Fregenbaum, J.M., Erbst, K. &

Newell, M.L. (2007).Continued very high prevalence of HIV infection in rural KwaZulu-Natal, South Africa: A populationbased longitudinal study. AIDS,21: 1467–72.

Wienke, A. (2011). Frailty Models in Survival data. Chapman and Hall, New York, 1 edition.

ZABA B, BOERMA T, PISANI E & BAPTIESTE N. 2002. Estimation of levels and trends in age at first sex from surveys using survival analysis. *Working Papers* [Online].

Appendix

Appendix A: Univariable Analysis of AFT Model using various base line Distributions

A. Weibull Uni-variable AFT Model for time to age at First Marriage and Sexual intercourse data

Variables	Age at First Marriage			Age at First Sexual Intercourse		
	Coef	ϕ (95%CI)	p-value	Coef	ϕ (95%CI)	p-value
Region						
Tigray						
Afar	Ref					
Amhara	-0.02	0.98(0.964,1.006)	0.167	0.03	1.03(1.006,1.048)	0.013
Oromiya	-0.07	0.93(0.916,0.953)	0.000	-0.05	0.96(0.938,0.974)	0.000
Somali	0.051	1.05(1.032,1.074)	0.000	0.09	1.09(1.076,1.117)	0.000
Benishangul-G	0.005	1.01(0.979,1.032	0.695	0.05	1.05(1.025,1.076)	0.000
SNNP	-0.04	0.96(0.942,0.985)	0.001	0.01	1.01(0.989,1.025)	0.379
Gambela	0.111	1.12(1.095,1.141)	0.000	0.15	1.17(-1.143,1.188)	0.000
Harari	-0.01	0.98(0.965,1.011)	0.304	0	1(0.981,1.025)	0.803
Dire Dawa	-0.04	0.96(0.936,0.992)	0.013	0.01	1.01(0.981,1.036)	0.573
	0.087	1.09(1.058,1.126)	0.000	0.13	1.14(1.108,1.175)	0.000
Women Education level						
No education	Ref					
Primary	0.13	1.13(1.119,1.14)	0.000	0.12	1.12(1.11,1.137)	0.000
Secondary	0.38	1.46(1.383,1.544)	0.000	0.34	1.41(1.338,1.478)	0.000
Higher	0.32	1.38(1.284,1.488)	0.000	0.22	1.24(1.168,1.323)	0.000
Mass media						
No	Ref					
Yes	0.001	1.001(0.99,1.01)	0.839	-0.01	0.99(0.983,1.004)	0.215
Wealth index						
Poor	Ref					
Middle	0.02	1.0221.008,1.036)	0.002	0.03	1.03(1.013,1.04)	0.000
Rich	0.04	1.0411.028,1.054)	0.000	0.04	1.04(1.028,1.053)	0.000
Women work status						
No	Ref					
Yes	0.014	1.01(1.002,1.026)	0.019	0.0136	1.01(1.003,1.025)	0.015

Religion						
Orthodox	Ref					
Muslim	0.012	1.01(1.000,1.024)	0.048	0.04	1.04(1.029,1.052)	0.000
Protestant	0.084	1.08(1.072,1.105)	0.000	0.11	1.11(1.098,1.13)	0.000
Others	0.023	1.02(0.992,1.055)	0.142	0.05	1.05(1.02,1.081)	0.001

Coef =coefficient, ϕ=acceleration factor, 95%CI= 95% confidence interval for acceleration factor, Ref = Reference

B. Log-logistic Univariable AFT Model for time to First Marriage and Sexual Intercourse data

Variables	Age at First Marriage			Age at First sexual intercourse		
	Coef	ϕ (95%CI)	p-value	Coef	ϕ (95%CI)	p-value
Region						
Tigray	Ref					
Afar	-0.01	0.99(0.973,1.006)	0.344	0.04	1.04(1.0186,1.0544)	0.000
Amhara	-0.13	0.88(0.865,0.894)	0.000	-0.1	0.94(0.9257,0.9549)	0.000
Oromiya	0.03	1.03(1.014,1.048)	0.000	0.07	1.08(1.0582,1.0918)	0.000
Somali	0.02	1.02(0.998,1.042)	0.073	0.06	1.07(1.0442,1.0871)	0.000
Benishangul-G	-0.05	0.95(0.930,0.966)	0.000	-0.01	0.99(0.9783,1.013)	0.610
SNNP	0.091	1.09(1.077,1.114)	0.000	0.13	1.14(1.1224,1.158)	0.000
Gambela	-0.02	0.98(0.963,1.002)	0.076	0.01	1.01(0.9881,1.024)	0.501
Harari	-0.01	0.99(0.966,1.014)	0.385	0.04	1.04(1.0124,1.059)	0.002
Dire Dawa	0.07	1.08(1.048,1.107)	0.000	0.12	1.13(1.099,1.1568)	0.000
Women Education level						
No education	Ref					
Primary	0.13	1.14(1.124,1.467)	0.000	0.12	1.12(1.114,1.135)	0.000
Secondary	0.34	1.41(1.352,1.463)	0.000	0.31	1.36(1.31,1.407)	0.000
Higher	0.34	1.41(1.335,1.485)	0.000	0.27	1.30(1.245,1.366)	0.000
Mass media						
No	Ref					
Yes	0.01	1.01(0.996,1.015)	0.257	0.002	0.62(0.993,1.011)	0.622
Wealth index						
Poor	Ref					
Middle	0.002	1.003(0.991,1.015)	0.685	0.01	1.01(0.995,1.018)	0.260
Rich	0.02	1.02(1.009,1.032)	0.0002	0.02	1.02(1.009,1.03)	0.000
Women work status						
No	Ref					
Yes	0.01	1.01(0.999,1.019)	0.084	0.01	1.01(0.998,1.017)	0.137

Religion						
Orthodox	Ref					
Muslim	0.04	1.04(1.030,1.052)	1.0411	0.0494	1.05(1.041,1.061)	0.000
Protestant	0.09	1.11(1.086,1.115)	1.1005	0.0984	1.10(1.09,1.117)	0.000
Others	0.06	1.06(1.032,1.089)	1.0604	0.0652	1.07(1.041,1.095)	0.000

Coef =coefficient, ϕ=acceleration factor, 95%CI= 95% confidence interval for acceleration factor, Ref = Reference

C. **Log-normal Univariable AFT Model for time to First Marriage and Sexual intercourse data**

Variables	Age at First Marriage			Age at First sexual intercourse		
	Coef	ϕ (95%CI)	p-value	Coef	ϕ (95%CI)	p-value
Region						
Tigray	Ref					
Afar	-0.004	0.99(0.9769,1.016)	0.713	0.04	1.04(1.018,1.056)	0.000
Amhara	-0.13	0.88(0.8649,0.895)	0.000	-0.06	0.94(0.924,0.955)	0.000
Oromiya	0.04	1.04(1.0203,1.057)	0.000	0.07	1.08(1.059,1.094)	0.000
Somali	0.03	1.03(1.0103,1.059)	0.005	0.07	1.08(1.052,1.099)	0.000
Benishangul-G	-0.05	0.95(0.9296,0.966)	0.000	-0.004	0.99(0.977,1.014)	0.636
SNNP	0.1	1.11(1.089,1.1290)	0.000	0.14	1.15(1.129,1.168)	0.000
Gambela	-0.01	0.99(0.968,1.0099)	0.296	0.01	1.01(0.988,1.028)	0.426
Harari	-0.005	0.99(0.969,1.022)	0.736	0.04	1.04(1.012,1.063)	0.003
Dire Dawa	0.07	1.08 (1.045,1.106)	0.000	0.11	1.12(1.091,1.151)	0.000
Women Education level						
No education	Ref					
Primary	0.132	1.14(1.129,1.153)	0.000	0.12	1.13(1.118,1.141)	0.000
Secondary	0.342	1.41 (1.353,1.464)	0.000	0.31	1.36(1.31,1.407)	0.000
Higher	0.342	1.41(1.329,1.492)	0.000	0.27	1.31(1.243,1.379)	0.000
Mass media						
No	Ref					
Yes	0.005	1.005(0.995,1.015)	0.342	0.001	1.001(0.991,1.01)	0.89
Wealth index						
Poor	Ref					
Middle	0.001	1.001(0.989,1.014)	0.825	0.01	1.01(0.994,1.018)	0.316
Rich	0.02	1.02 (1.009,1.032)	0.001	0.02	1.02(1.009,1.03)	0.000

Women work status						
No	Ref					
Yes	0.01	1.01(0.997,1.018)	0.192	0.01	1.01(0.996,1.015)	0.289
Religion						
Orthodox	Ref					
Muslim	0.05	1.05(1.039,1.061)	0.000	0.05	1.05(1.043,1.064)	0.000
Protestant	0.11	1.12(1.101,1.131)	0.000	0.11	1.11(1.098,1.126)	0.000
Others	0.06	1.06 (1.035,1.096)	0.000	0.06	1.07(1.039,1.095)	0.000

Coef =coefficient, ϕ=acceleration factor, 95%CI= 95% confidence interval for acceleration factor, Ref = Reference

Appendix B: Multivariable nalysis of AFT Model using various base line Distribution

A. Weibull Multi-variable AFT Model for time to First Marriage and Sexual Intercourse data

Variables	Age at First Marriage			Age at First sexual intercourse		
	Coef	ϕ (95%CI)	p-value	Coef	ϕ (95%CI)	p-value
Region						
Tigray	Ref					
Afar	0.03	1.03(0.999,1.057)	0.056	0.06	1.06(1.036,1.092)0.	0.000
Amhara	-0.06	0.94(0.922,0.959)	0.000	-0.04	97(0.947,0.984)	0.000
Oromiya	0.05	1.05(1.029,1.080)	0.000	0.08	1.09(1.065,0.109)	0.000
Somali	0.05	1.05(1.0168,1.08)	0.003	0.08	1.09(1.057,1.122)	0.000
Benishangul-G	-0.03	0.97(0.948,0.998)	0.036	0.01	1.01(0.990,1.039)	0.248
SNNP	0.10	1.11(1.081,1.139)	0.000	0.14	1.15(1.118,1.174)	0.000
Gambela	-0.02	0.98(0.948,1.004)	0.097	-0.01	0.97(0.959,1.013)	0.302
Harari	-0.01	0.99(0.952,1.022)	0.45	0.02	1.02(0.989,1.056)	0.202
Dire Dawa	0.12	1.12(1.083,1.164)	0.000	0.15	1.16(1.125,1.205)	0.000
Women Education level						
No education	Ref					
Primary	0.12	(1.1126,1.1426)	0.000	0.11	1.12(1.104,1.131)	0.000
Secondary	0.38	(1.3794,1.5393)	0.000	0.34	1.40(1.335,1.472)	0.000
Higher	0.36	(1.3262,1.5365)	0.000	0.25	1.28(1.204,1.361)	0.000
Wealth index						
Poor	Ref					
Middle	0.003	1.003(0.989,1.017)	0.704	0.01	1.01(0.995,1.022)	0.231
Rich	-0.01	0.98(0.973,1.0001)	0.512	-0.01	0.99(0.979,1.005)	0.237

Womenwork status						
No	Ref					
Yes	-0.01	0.98(0.975,0.994)	0.027			
Religion						
Orthodox	Ref					
Muslim	-0.004	0.99(0.978,1.015)	0.675	0.01	1.01(0.990,1.025)	0.384
Protestant	0.01	1.01(0.988,1.031)	0.383	0.03	1.03(1.007,1.047)	0.009
Others	-0.01	0.99(0.959,1.026)	0.621	0.01	1.01(0.981,1.04)	0.449

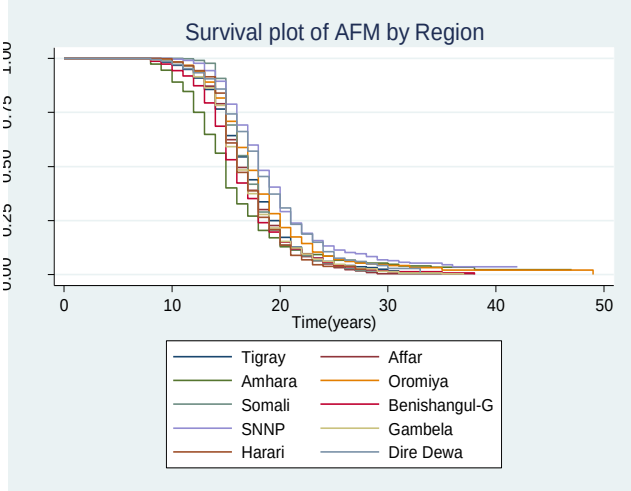
Coef =coefficient, ϕ=acceleration factor, 95%CI= 95% confidence interval for acceleration factor, Ref = Reference

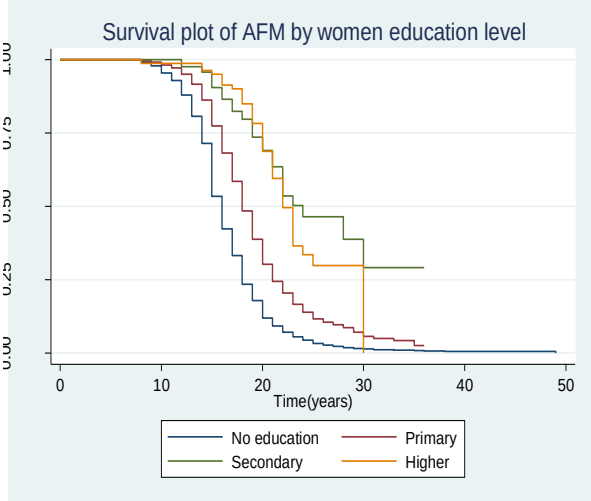
B. Log-normal Multi-variable AFT Model for time to First Marriage and Sexual Intercourse data

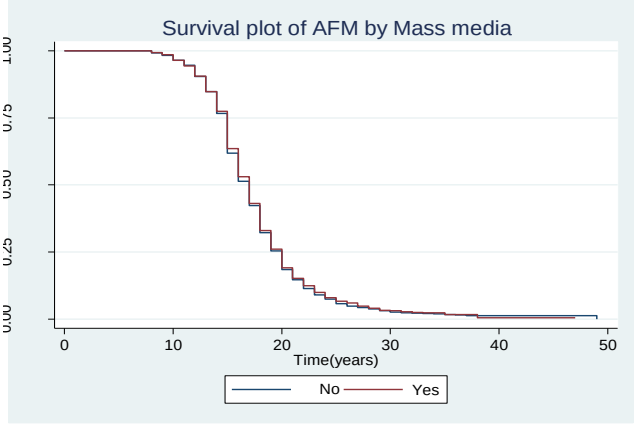
Variables	Age at First Marriage			Age at First sexual intercourse		
	coef	ϕ (95%CI)	p-value	Coef	ϕ (95%CI)	p-value
Region						
Tigray	Ref					
Afar	0.02	1.02(0.991,1.040)	0.208	0.06	1.06(1.039,1.086)	0.000
Amhara	-0.12	0.89(0.874,0.905)	0.000	-0.05	0.95(0.935,0.965)	0.000
Oromiya	0.02	1.02(1.003,1.045)	0.026	0.06	1.06(1.044,1.084)	0.000
Somali	0.05	1.06(1.028,1.086)	0.000	0.09	1.10(1.075,1.131)	0.000
Benishangul-G	-0.06	0.94(0.922,0.963)	0.000	-0.01	0.99(0.973,1.014)	0.503
SNNP	0.08	1.08(1.059,1.108)	0.000	0.11	1.12(1.099,1.145)	0.000
Gambela	-0.05	0.95(0.925,0.972)	0.000	-0.03	0.97(0.95,0.989)	0.004
Harari	0.00	1(0.970,1.031)	0.998	0.05	1.05(1.019,1.077)	0.001
Dire Dawa	0.08	1.09(1.052,1.120)	0.000	0.13	1.14(1.104,1.171)	0.000
WomenEducation level						
No education	Ref					
Primary	0.13	1.14(1.127,1.151)	0.000	0.12	1.13(1.121,1.143)	0.000
Secondary	0.35	1.42(1.368,1.475)	0.000	0.32	1.38(1.331,1.426)	0.000
Higher	0.38	1.46(1.376,1.539)	0.000	0.30	1.35(1.286,1.423)	0.000
Wealth index						
Poor	Ref					
Middle	-0.01	0.99(0.981,1.006)	0.277	-0.003	0.99(0.985,1.008)	0.509
Rich	-0.03	0.97(0.964,0.986)	0.000	-0.02	0.98(0.967,0.988)	0.000
Womenwork status						
No	Ref					
Yes	-0.01	0.99(0.983,1.001)	0.204			
Religion						
Orthodox	Ref					
Muslim	0.02	1.02(1.003,1.034)	0.021	0.01	1.01(0.997,1.026)	0.109
Protestant	0.03	1.03(1.009,1.046)	0.003	0.03	1.03(1.012,1.046)	0.001
Others	0.02	1.02(0.987,1.047)	0.283	0.02	1.02(0.992,1.048)	0.158

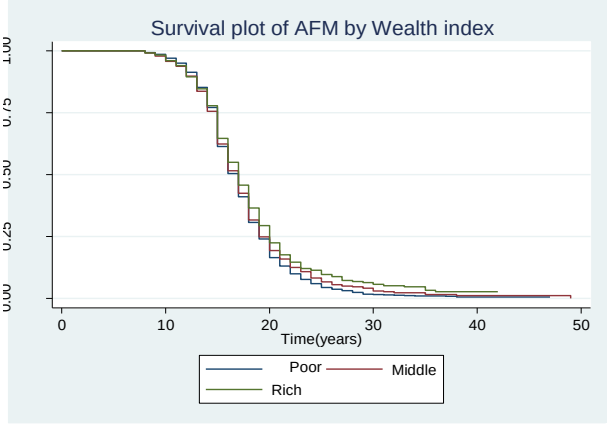
Coef =coefficient, ϕ=acceleration factor, 95%CI= 95% confidence interval for acceleration factor, Ref = Reference

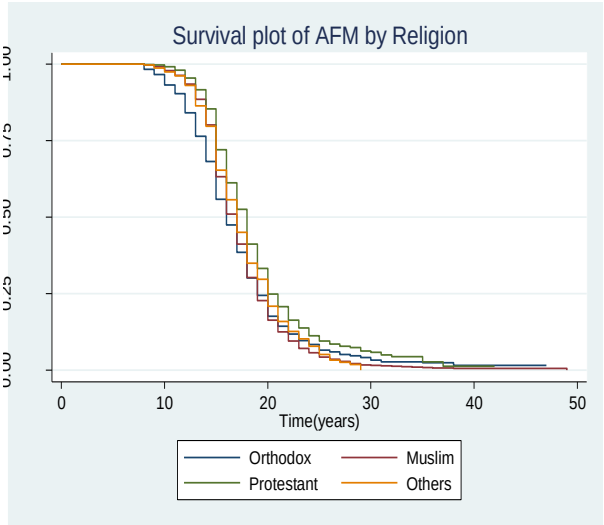
Appendix C: K-M Plots of Survival of Time to Age at First Marriage and Age at First Sexual Inter course by Different Categorical Group of Covariates

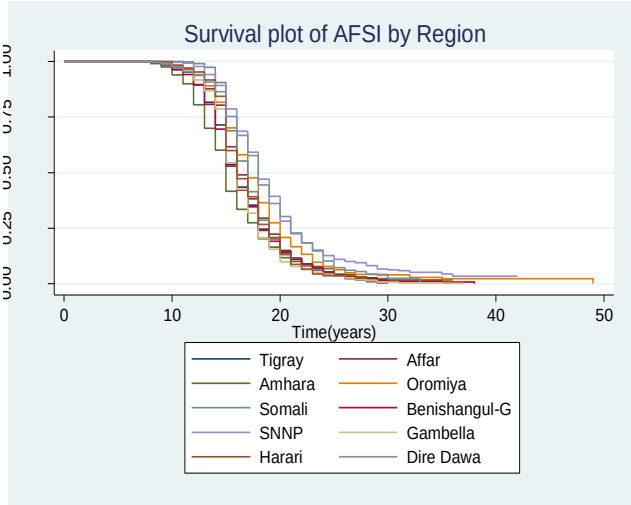


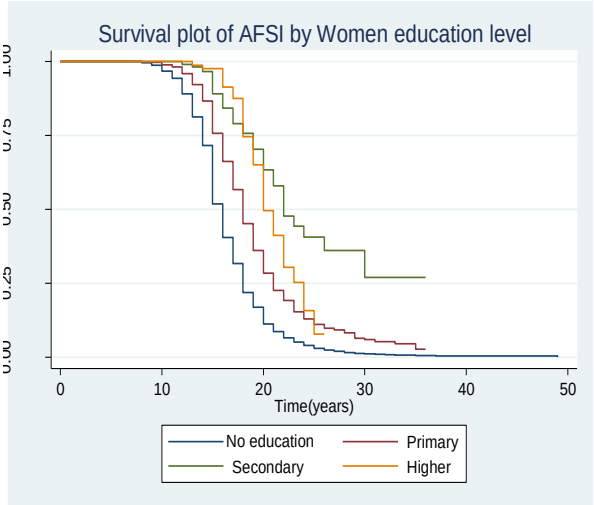


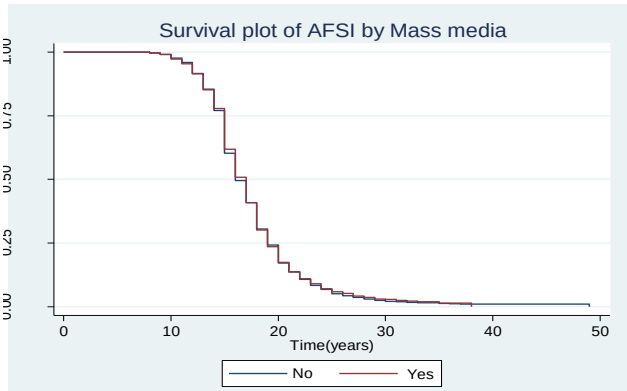


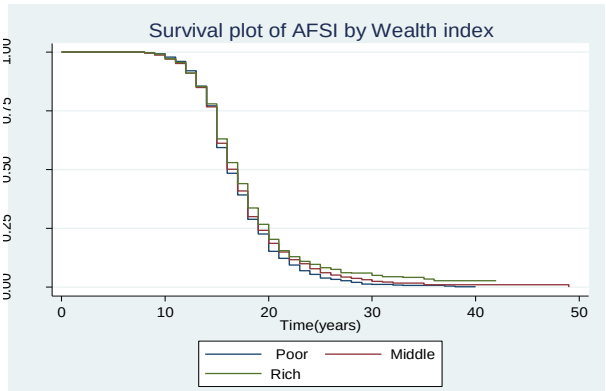


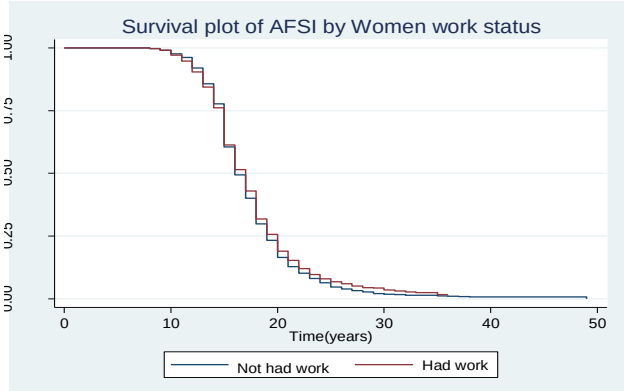


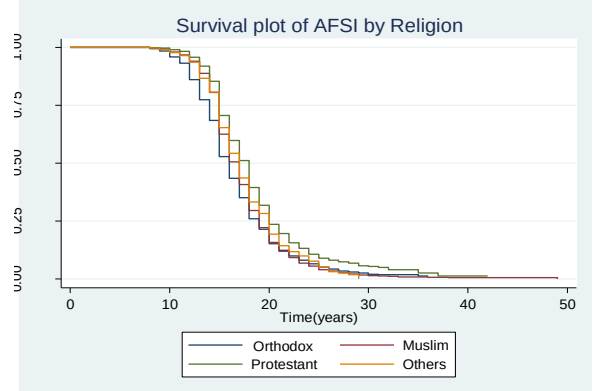

Survival plot of AFSI by Region
0.00
0.25
0.50
0.75
1.00
0
10
20
30
40
50
Time(years)
Tigray
Affar
Amhara
Oromiya
Somali
Benishangul-G
SNNP
Gambella
Harari
Dire Dawa


Survival plot of AFSI by Women education level
0.00
0.25
0.50
0.75
1.00
0
10
20
30
40
50
Time(years)
No education
Primary
Secondary
Higher


Survival plot of AFSI by Mass media
0.00
0.25
0.50
0.75
1.00
0
10
20
30
40
50
Time(years)
No
Yes


Survival plot of AFSI by Wealth index
0.00
0.25
0.50
0.75
1.00
0
10
20
30
40
50
Time(years)
Poor
Middle
Rich


Survival plot of AFSI by Women work status
0.00
0.25
0.50
0.75
1.00
0
10
20
30
40
50
Time(years)
Not had work
Had work


Survival plot of AFSI by Religion
0.00
0.25
0.50
0.75
1.00
0
10
20
30
40
50
Time(years)
Orthodox
Muslim
Protestant
Others