Modeling the Determinant of Time to Age at First Marriage among Women in Ethiopia using Cox models with mixed effects

Molalign Gualu Gobena (✉ molaligngualu@gmail.com)

Assosa University <https://orcid.org/0000-0002-5281-9806>

Yebelay Berelle Berhan

Debre Markos University

Research

Keywords: Grouped Data, Heterogeneity, Random effect, Shared frailty, Survival

Posted Date: July 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-684041/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Reproductive Health on January 31st, 2022.
See the published version at <https://doi.org/10.1186/s12978-022-01339-4>.

Abstract

Background

Time to age at first marriage of women is the duration of time until the age at which they started living with their first partner. Time to age at first marriage is widely considered a proxy indicator for the age at which women begin to be exposed to the risks inherent in sexual activity. The purpose is to model the determinant of time to age at first marriage among women in Ethiopia using Cox models with mixed effects.

Methods

The 2016 EDHS sample was selected using a two-stage cluster design. The data set in this study were obtained from the Demography and Health survey conducted in Ethiopia in 2016. In this study, we used Cox models with mixed effects.

Results

Of all 15,683 women aged 15–49, 11,405(72.72%) were married and the median & mean age at first marriage for women living in Ethiopia were 17 years and 18 years respectively.

Conclusions

Based on the AIC criterion Cox frailty survival model (log-normal frailty distribution) was selected among Cox models with mixed effects as a better fit for the dataset. Analyzing based on the selected model showed that residence of women, educational level of women, head occupation, work status of women & head educational level were the most significant factors and religion of women, access to media of women & wealth index of a household were not significant factors at 5% level of significance. The clustering effect was significant and hence there was heterogeneity among the regions on age at first marriage.

1. Background

Time to age at first marriage of women is the duration of time until the age at which they started living with their first partner. Time to age at first marriage is widely considered a proxy indicator for the age at which women begin to be exposed to the risks inherent in sexual activity. A comparison of the median age at first intercourse with the median age at first marriage can be used as a measure of whether women engage in sex before marriage. The median age at first sexual intercourse for women age 25–49 years is 16.6 years, which is very close to the median age at first marriage of 16.5 years. This suggests that Ethiopian women generally begin sexual intercourse at the time of their first marriage. In Ethiopia

marriage marks the point in a woman's life when childbearing becomes socially acceptable. Age at first marriage has a major effect on childbearing because women who marry early have on average a longer period of exposure to the risk of pregnancy and give birth to a greater number of children over their lifetimes (1). African women are more likely to marry earlier than other continents women, which causes high fertility due to their long period of exposure to the risk of pregnancy. Even though, Sub-Sahara Africa accounts for the highest rate of age at first marriage among countries in the Africa continent, comparably the case is very worst in Ethiopia (2).

Very early age at first marriage and premarital first sex are associated with marital instability and divorce, multiple partners; poverty, and subsequent drift into prostitution or paid domestic work (3). The age at first marriage may also influence population growth, labor supply, consumption, wage rates, mortality, migration, and to some extent fertility (4). Women who marry early will have, on average; a longer period of exposure to the risk of pregnancy, often leading to higher completed fertility. Variation in the age of entry into marriage helps explain differences in fertility across populations and also helps explain trends in fertility within individual populations over time (5). Those aforementioned are one reason why age at first marriage is our source of concern in Ethiopia. Even though, there is some study about the determinant of time to age at first marriage in Ethiopia using different models such as logistic regression model, cox proportional hazard models, etc. There is no study about the determinant of time to age at first marriage in Ethiopia using advanced models like Cox Models with Mixed-effects. Using these above two models for this study leading to misleading results because the logistic regression model does not consider censoring observation and cox proportional hazard models consider identical & independent distribute samples.

As a result, here we want to use Cox models with mixed effects since this model permits the analysts to account for the loss of independence that arises from the clustering of subjects in higher-level units (6). The specific objectives of this study are to:- Distinguish significant factors or covariates that are related to time-to-age at first marriage and estimate the variance of the random effect distribution for the data set. The study helps to indicate relevant solutions for women marriage-related problems in Ethiopia after analyzing the data set and the study will also add to the existing literature on the determinant of time to age at first marriage, that is, it provides input for further study in Ethiopia.

2. Methodology

2.1 Data Source

The data set in this study was obtained from the EDHS conducted in 2016, which was the fourth comprehensive survey conducted as part of the worldwide DHS project (7).

2.2 Sample Design

The 2016 EDHS sample was selected using a two-stage cluster design. Census enumeration areas (EAs) were the sampling units for the first stage and women in household are the sampling unit for the second

stage. The survey was implemented by (CSA, Ethiopia) from January 18, 2016, to June 27, 2016. The survey selected 645 EAs, 202 in urban areas, and 443 in rural areas. The survey was conducted in 16,650 residential households, 5,232 in urban areas, and 11,418 in rural areas. The sample was expected to generate an estimated 16,663 completed interviews with women age 15–49, 5,514 in urban areas and 11,149 in rural areas. Women whose current ages are 15–49 years are included in the survey (7). After a certain rearrangement and reorganization, the total number of women with complete information became 15,683.

2.3 Variable in the Study

The dependent variable is the time to age at first marriage. It is measured as the length of time from birth until the age at first marriage which is measured in years. The independent variables considered in this study are the respondent's work status, religion, type of residence, head education level, women education level, head occupations, access to media, and wealth index.

2.4 Method of Data Analysis

2.4.1 Survival Analysis

Survival analysis consists of studies of the survival time of a subject (usually measured in days, weeks, months, or years), which is the time that elapses between the baseline and the moment an adverse event occurs, or the subject drops out of the trial. The survival times for subjects who dropped out of the trial are right-censored. The survival times of the subjects who remain in the trial until it ends are censored *as well*. In what follows, each uncensored observation is termed "death," regardless of whether death or a different adverse event has occurred. Denote by T the random variable representing the survival time of a subject. Let $f(t)$, $t \geq 0$, denote the probability density function (pdf) of T , and let $F(t) = P(T \leq t)$, $t \geq 0$, be the cumulative distribution function (CDF) of T . The distribution of T is called the survival time distribution. The survival function, $S(t)$, defined as the probability that a subject survives up to time t :

$$S(t) = P(T > t) = \int_t^{\infty} f(x) dx = 1 - F(t), t \geq 0 \dots \dots \dots (2.1)$$

(8).

2.4.2 Median Survival Times

The median survival times to be the smallest value of t for which ≤ 0.5 , that is, the time t where it jumps from a value greater than 0.5 to a value less than or equal to 0.5 (8).

2.4.3 Cox models with mixed effects

Multilevel or grouped data like individuals are nested within families, & families are nested within neighborhoods that are common across a wide range of fields of study. This study also encountered such kinds of data. For instance, women aged 15–49 are nested within the region. As a result, it is two-level data. The inclusion of random effects into a Cox proportional hazards model shares many

similarities with methods for the analysis of multilevel data with continuous, binary, or count outcomes. Cox proportional hazards model is enhanced through the incorporation of random effect terms to account for within-cluster homogeneity in outcomes. Applying the Cox proportional hazards regression to such grouped survival times leads to biased tests of statistical significance (9,10). Moreover, cox's model needs identically and independently distributed samples. Cox regression models with mixed effects do not assume as the observations are independent and allowed to apply for grouped data since it is one of the statistical models for multilevel survival analysis. The Cox regression model with mixed effects is said to be a frailty model when it is applied for two-level data.

Early frailty models incorporated subject-specific random effects to account for unmeasured subject characteristics that influenced the hazard of the occurrence of the outcome. These models were then extended to models that incorporate cluster-specific random effects to account for within-cluster homogeneity in outcomes. These models have been described as shared frailty models because the same random effect is shared by all subjects within the same cluster. As a result, the Cox regression model with mixed effects is said to be a shared frailty model when it is a model that incorporates cluster-specific random effects to account for within-cluster homogeneity in outcomes and particularly when it is applied for two-level data. For instance, in this study women within the same region share the same random effect concerning marriage (6). However, here we have used the special case of the Cox regression model with mixed effects what we call the shared frailty model to account for within region homogeneity in the marriage of women. When random effects are incorporated in the Cox model, these random effects denote increased or decreased hazard for distinct classes. Suppose individuals are nested in one of G groups or clusters. A mixed-effects Cox regression model can be formulated as:-

$$h_i(t) = h_0(t) \exp(x_i \beta + \alpha_j) \dots\dots\dots (2.2)$$

Where α_j denotes the random effects associated with the j^{th} cluster. Rabe-Hesketh (11) use the term 'shared frailty' to denote the exponential of the random effect: $\exp(\alpha_j)$. The random effect can be thought of as a random intercept that modifies the linear predictor, while the shared frailty term has a multiplicative effect on the baseline hazard function:

$$h_i(t) = h_0(t) \exp(\alpha_j) \exp(x_i \beta + \alpha_j) \dots\dots\dots (2.3)$$

Cox regression models with mixed effects are characterized by the distribution of the shared frailty terms. Different distributions have been proposed for the distribution of the shared frailty terms, including the gamma distribution, the log-normal distribution (the frailty terms will have a log-normal distribution while the random effects will have a normal distribution), positive stable frailty distributions, and power variance function distributions. The first two appear to be the most commonly used. In the gamma frailty model, the cluster-specific random effects are distributed as the logarithms of independent, identically distributed gamma random variables, having variance θ . In the log-normal frailty model, the cluster-specific random effects are distributed as the natural logarithms of independent, identically distributed

normal random variables, having variance θ . Generally, mixed-effects Cox regression models are used to model survival data when there are repeated measures on an individual, individuals nested within some other hierarchy, or some other reason to have both fixed and random effects (12,13).

2.5 Comparisons of Models

Even though there are so many model selection criteria, AIC is the most familiar model selection criteria (14). Therefore, here we used AIC criteria to compare two different Cox shared frailty models [i.e Cox frailty survival model (log-normal frailty distribution) and Cox frailty survival model (gamma frailty distribution)]. Cox shared frailty model with the least AIC value is taken as the best-fitted model for the data set.

3. Results

3.1 Descriptive Statistics

Of all 15,683 women aged 15–49, 11,405(72.72%) were married and the median & mean age at first marriage for women living in Ethiopia were 17 years and 18 years respectively, while the minimum and maximum age at first marriage observed were 10 years and 50 years respectively.

3.2 Multivariable Survival Analysis for Cox models with mixed effects

In this study, we were done multivariable survival analysis using Cox models with mixed effects. We were used the AIC criterion to compare models under Cox models with mixed effects. Accordingly, the Cox frailty survival model (log-normal frailty distribution) was selected as a better fit for the dataset compared to the Cox frailty survival model (gamma frailty distribution). Their AIC values are 193779.4 & 193779.7 respectively.

Table 1
Multivariable analysis using the Cox frailty survival model (log-normal frailty distribution).

Covariate	Coeff	St.err	P-value	ϕ
Residence	1	0.0307	1.1e-04	0.8880
Urban(ref)	-0.1188			
Rural				
Education Level	1	0.0524	4.3e-08	1.3325
No Education(ref)	0.2871	0.0495	0.000	2.2953
Primary	0.8309	0.0520	0.000	2.8354
Secondary	1.0422			
Higher				
Religion	1	0.1258	6.8e-01	0.9498
Catholic (ref)	-0.0516	0.1255	5.1e-01	0.9200
Orthodox	-0.0834	0.1256	9.1e-01	0.9854
Protestant	-0.0148	0.1506	5.4e-01	0.9113
Muslim	-0.0929			
Others				
Wealth Index	-0.0338	0.0353	3.4e-01	0.9668
Middle (ref)	-0.0110	0.0323	5.4e-01	0.9802
Poorer	-0.0363	0.0353	3.0e-01	0.9644
Poorest	-0.0454	0.0337	1.8e-01	0.9556
Richer				
Richest				
Head Education	1.5538	0.0326	0.000	4.7293
No Education(ref)	1.5684	0.0387	0.000	4.7990
Primary	1.6063	0.0430	0.000	4.9841
Secondary				
Higher				

Covariate	Coeff	St.err	P-value	ϕ
Work Status	0.1239	0.0208	2.6e-09	1.1319
No(ref)				
Yes				
Access to media	-0.0273	0.0239	2.5e-01	0.9730
Yes (ref)				
No				
Head Occupation	-0.0485	0.0255	5.7e-02	0.9526
Professional (ref)	-0.0951	0.0400	1.7e-02	0.9092
Agriculturalist	-0.0692	0.0373	6.4e-02	0.9331
Laborers	-0.0228	0.0431	6.0e-01	0.9774
Business				
Others				
Random effect				
Group name	Variance	P-value		
Region	0.9972	0.000		

Analysis based on Cox frailty survival model (log-normal frailty distribution) showed that place of residence of women, one category of head occupation, education level of women, work status of women, and head education was significant at 5% level of significance (since their coefficient p-value < 5%). In contrast religion of women, access to media of women, and wealth index of a household were not significant at the 5% level of significance (since their coefficient p-value > 5%).

An odds ratio greater than 1 indicates that women with that category are more likely to extend their age at first marriage than women without that category. Therefore women who attend primary school ($\phi=1.3325$), secondary school ($\phi=2.2953$), and higher education ($\phi=2.8354$) are more likely to extend their age at first marriage by a factor of 1.3325, 2.2953 & 2.8354 respectively than illiterate women (no education). Similarly, those women having a head that attends primary school, secondary school and higher education are more likely to extend their age at first marriage by a factor of 4.7293, 4.7990, and 4.9841 respectively from those having an illiterate head (no education). In the same way, women who had work are 1.1319 times more likely to extend their age at first marriage than those who haven't work (reference category).

Categories of significant covariates having odds ratio less than 1 imply that women characterized by those categories of the same covariate are less likely to extend their age at first marriage than those

women who are characterized by the reference category of the same covariates. For instance, Women residing in a rural area of Ethiopia have married early than those residing in an urban area (ref) of Ethiopia ($\phi=0.8880$). Thus, those women who have laborer heads ($\phi=0.9435$) marry early than women who have a professional head. The variability (heterogeneity) among regions in Ethiopia with regards to the age at first marriage for women which is estimated by the Cox frailty survival model (log-normal frailty distribution) was 0.9972. The test shows that there are a significant variation regards to the age at first marriage of women among the region of Ethiopia.

4 Discussion

Analyses based on the Cox frailty survival model (log-normal frailty distribution) explain that women who are living in a rural area are more likely to marry early than in an urban area. This result of ours coincides with a study conducted in Ethiopia by Tessema B (15). Similarly, according to Tessema B (15) women who had work are more likely to marry later than those who have not. This also coincides with our study. This implies that the economic advantage of women had a great role in delaying the age at first marriage in Ethiopian women. A study in Nigeria by Adedokun OA (16) shows that the increment of mean age at first marriage when the duration of schooling increases. This coincides with our finding since our study revealed that better-educated women are more likely to extend their age at first marriage than illiterate women (no education). Head education is also positively related to age at first marriage. i.e those women having better-educated head had extended age at first marriage (17). Similarly, a study by Bates LM (18) suggests that the increment of the level of education for head gives to raise awareness for women and as a result of it leads to higher age at first marriage. These also support our study. According to Makhete FM (19) wealth index of household improve age at first marriage (i.e women favor later marriage). Under this study religion also other significant factors. The finding of this study contradict our conclusion since access to media, wealth index of household, a religion of women, and some categories of head occupation were not significant factors at a 5% level of significance.

5. Conclusion

Based on the AIC criterion Cox frailty survival model (log-normal frailty distribution) was selected among Cox models with mixed effects as a better fit for the dataset. As a result, analyzing using the Cox frailty survival model (log-normal frailty distribution) showed that residence of women, educational level of women, head occupation, work status of women, and head educational level were the most significant factors for the time-to-age at first marriage. Women residing in the urban part of Ethiopia had an extended age at first marriage as compared to those residing in the rural part of Ethiopia. Concerning the educational level of women, women having better education had an extended age at first marriage than illiterate women. Similarly, women having Head who achieved higher education had an extended age at first marriage compared to women having an illiterate head. Concerning, the work status of women, those having work marry later than those who haven't work. This study also showed that age at first marriage for women living in Ethiopia varies across the region.

This study also revealed that, of all 15,683 women age 15–49, 11,405(72.72%) were married and the median & mean age at first marriage for women living in Ethiopia were 17 years and 18 years respectively, while the minimum and maximum age at first marriage observed were 10 years and 50 years respectively.

Abbreviations

AIC: Akaike Information Criteria; **CSA:** Central Statistical Agency; **EAs:** Enumeration Areas; **DHS:** Demographic and Health Survey; **EDHS:** Ethiopian Demographic and Health Survey

Declarations

Acknowledgments

We have the deepest gratitude to the CSA of Ethiopia for making available the data used in this study.

Author's Contributions

We have a contribution to the study throughout writing the introduction, methodology, result, discussion, conclusion, and recommendation parts of the study. Additionally, we have reviewed the entire part of the study.

Funding

Not Applicable

Availability of Data and Material

The data that support the finding of this study is available from <http://www.measuredhs.com>.

Ethics Approval and Consent to Participate

It is impossible to get the participants in this study because all personally identifiable information was removed in this dataset. But, permission to use the data was obtained from the Ethiopian Central Statistical Agency. Therefore, additional ethical approval may not be desired.

Consent for Publication

Not Applicable

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Demographic N. Health Survey 2011. Addis Ababa, Ethiop. 2012;
2. Bayisenge J. Early marriage as a barrier to girl's education. Retrieved on. 2010;1(10):2010.
3. Lindstrom DP, Kiros G-E, Hogan DP. Transition into first intercourse, marriage, and childbearing among Ethiopian women. *Genus*. 2009;65(2):45.
4. Mensch BS, Singh S, Casterline JB. Trends in the timing of first marriage among men and women in the developing world. *Chang transitions to adulthood Dev Ctries Sel Stud*. 2005;118–71.
5. EZEH AC, Dadoo FN-A. Institutional change and the African fertility transition: the case of Kenya. *Genus*. 2001;135–64.
6. Crowther MJ, Look MP, Riley RD. Multilevel mixed effects parametric survival models using adaptive Gauss–Hermite quadrature with application to recurrent events and individual participant data meta-analysis. *Stat Med*. 2014;33(22):3844–58.
7. Csa I. Central Statistical Agency (CSA)[Ethiopia] and ICF. Ethiopia Demographic and Health Survey, Addis Ababa. Ethiopia and Calverton. 2016.
8. Collett D. Modelling survival data in medical research. CRC press; 2015.
9. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat*. 1982;1100–20.
10. Prentice RL, Williams BJ, Peterson A V. On the regression analysis of multivariate failure time data. *Biometrika*. 1981;68(2):373–9.
11. Rabe-Hesketh S, Skrondal A. Multilevel and longitudinal modeling using Stata. STATA press; 2008.
12. Hougaard P. Analysis of multivariate survival data. Springer Science & Business Media; 2012.
13. Wienke A. Frailty models in survival analysis. CRC press; 2010.
14. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19(6):716–23.
15. Tessema B, Ayalew S, Mohammed K. Modeling the determinants of time-to-age at first marriage in ethiopian women: A comparison of various parametric shared frailty models. *Sci J with Public Heal*. 2015;3(5):707–18.
16. Adedokun OA. Shifting patterns of age at first marriage and implications for family size in metropolitan Lagos (Nigeria). 1999;

17. Michael RT, Tuma NB. Entry into marriage and parenthood by young men and women: The influence of family background. *Demography*. 1985;22(4):515–44.
18. Bates LM, Maselko J, Schuler SR. Women's education and the timing of marriage and childbearing in the next generation: evidence from rural Bangladesh. *Stud Fam Plann*. 2007;38(2):101–12.
19. Makhete FM. Determinants of Age at first marriage in Kenya: A comparative analysis between men and women. 2008.