

Modeling the Determinants of Time-to-age at First Marriage in Ethiopian Women: A Comparison of Various Parametric Shared Frailty Models

Bedasa Tessema¹, Salie Ayalew², Kasim Mohammed^{2, *}

¹Department of Statistics, College of Natural & Computational Sciences, Drie-dawa University, Drie-dawa, Ethiopia

²Department of Statistics, College of Natural & Computational Sciences, University of Gondar, Gondar, Ethiopia

Email address:

t_bedasa@yahoo.com (Bedasa T.), salie.ayalew@gmail.com (Salie A.), kasim.fti@gmail.com (Kasim M.)

To cite this article:

Bedasa Tessema, Salie Ayalew, Kasim Mohammed. Modeling the Determinants of Time-to-age at First Marriage in Ethiopian Women: A Comparison of Various Parametric Shared Frailty Models. *Science Journal of Public Health*. Vol. 3, No. 5, 2015, pp. 707-718.

doi: 10.11648/j.sjph.20150305.27

Abstract: Marriage is an important part of human life and age at first marriage is the age at which individuals get married. This varies across communities and individuals in different country. Ethiopia is one of the Sub-Saharan Africa in which highest at early marriage and a small number of delayed marriages are occurred. Survival analysis is a statistical method for data analysis where the outcome variable of interest is the time to the occurrence of an event. Frailty model is an extension of Cox's proportional hazard model in which the hazard function depends upon an unobservable random quantity, the so-called frailty. Regional states of the women were used as a clustering effect in all frailty models. The study aimed to model the determinants of time-to-age at first marriage in Ethiopia. The data source for the analysis was the 2011 EDHS data collected during September 2010 through January 2011 from which the survival information of 12208 woman on age at first marriage. The gamma and inverse Gaussian shared frailty with exponential, Weibull and log-logistic baseline models was employed to analyze risk factors associated with age at first marriage using socio-economic and demographic factors. All the fitted models were compared by using AIC. Out of the total, about 69.3% of women were married and 30.7% were not married at different age of marriage. The median of age at first marriage was 17 years. The log-logistic with inverse Gaussian shared frailty model had minimum value of AIC when compared with other models for age at first marriage dataset. The clustering effect was significant for modeling the determinants of time-to-age at first marriage dataset. Based on the result of log-logistic-inverse Gaussian shared frailty model, women educational level, head/parents occupation, place of residence, educational level of head/parents, access to media and respondent work status were found to be the most significant determinants of age at first marriage. The estimated acceleration factor for the group of women's who had secondary and higher educational level were highly prolonged age at first marriage by the factor of $\phi=1.0796$ and $\phi=1.1497$ respectively. The log-logistic with inverse Gaussian shared frailty model described age at first marriage dataset better than other models and there was heterogeneity between the regions on age at first marriage. Improving girls and young women access to education was an important avenue for rising women's age at first marriage and for empowering women.

Keywords: Time-to-age at First Marriage, Risk Factors, Comparison of Models

1. Introduction

Marriage is an important family institution for the individual and the society at large. For the individual, it is a significant and memorable event in one's life cycle as well as the most important foundation in the family formation process. It is also a rite of passage that marks the beginning of an individual's separation from the parental unit, even if generations continue to be socially and economically

interdependent. For the society as a whole, it 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. In addition, marriage marks the beginning to an end of the transition to adulthood as the individual separates from the parental home, even if generations continue to be socially and economically

interdependent through the extended family. A marriage is a legally recognized union between a man and a woman in which they are united sexually; cooperate economically, and may have children through birth or adoption (Ikamari, 2005).

Age at marriage is the age at which individuals get married and this varies across communities and individuals. Marriage a time may be wanted or unwanted at a particular time. The term "age at early marriage" is used to refer both formal marriages and informal unions in which a girl lives with a partner as if married before age of 18 (UNICEF, 2005; Forum on Marriage and the rights of women and girls, 2001). According to UNFPA (2006), early marriage, also known as Child marriage, is defined as "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". Age at 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. Age at early marriage is common in much of the developing world, adolescent and age at child marriage continues to be a strong social norm, particularly for girls. It is associated with early childbearing, in most cases particularly in the developing world; the main purpose of marriage is to have children (UNFPA, 2006).

According to Demographic and Health Surveys (DHS), which provide much of the current country-level child marriage data, age at child marriage is most common in the world's poorest countries. The highest rates are in sub-Saharan Africa and South Asia as well as parts of Latin America and the Caribbean (NRC/IOM, 2005). A UNICEF study found that 48 percent of women were married before 15 years and 24 were married before 18 years in South Asia. The prevalence of age at early marriage is 42 percent in Africa (UNICEF, 2005) and more than 60 percent in some parts of East and West Africa (IPPF and UNFPA, 2006). In Latin America and the Caribbean, prevalence is 29 percent, though some individual countries have much higher rates of age at early marriage (UNICEF, 2005). Also age at child marriage is common in the Middle East, where nearly half of girls younger than 18 in Yemen and Palestine are married (IPPF and UNFPA, 2006). In sub-Saharan Africa, for example, 21 of 30 countries have seen an increase in the national age at marriage over the past several decades (Westoff, 2003). This increase in the age at marriage is occurring slowly and unevenly within countries, however, and many girls are missed by this trend. According to (UNICEF's, 2011) figures, 66 percent of Bangladesh girls are married before the age of 18 and approximately a third of women were married by the age of 15 ; although the legal age at first marriage for females in Bangladesh is 18 years.

The highest rates of child marriage are found in West Africa, in countries such as Niger, Chad, and Mali. However, in East Africa, the numbers of girls married in countries such as Ethiopia, Zambia, and Tanzania is also substantial. In rural Tanzania, median age at marriage is 18.5. The Demographic and Health Survey (DHS) for 1995 to 2003 shows that in

Niger, 47 percent of women aged between 20 and 24 were married before the age of 15, and 87 percent before the age of 18, a total of 53 percent had also had a child before the age of 18. Bayisenge (2010) observed that African women in general marry at a much earlier age than their non-African counterparts, leading to early pregnancies. In average, age at first marriage is relatively high, compared with developed countries and many other developing countries.

Ethiopia has one of the highest rates of age at early marriage in Sub-Saharan Africa. 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 recent study conducted in two woredas 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).

Age before 18 years marriage stands in a direct conflict with the objectives of the Millennium Development Goals (MDGs) (Mathur *et al.*, 2003). It threatens the achievement of MDGs such as eradicating extreme poverty and hunger, achieving universal primary education, promoting gender equality and empowering women, reducing child mortality, improving maternal health and combating HIV/AIDs, malaria and other diseases (UN, 2007).

In this study, we used shared frailty models by assuming that marriage within the same cluster (region) shares similar risk factors, which could be taken care of the frailty term at regional level. This model is a conditional independence model where the frailty is common to all individuals in a cluster and therefore responsible for creating dependence between event times. Parametric frailty models are used to investigate the relationship between different potential covariates and time-to-age at first marriage for clustered survival data with a random right censoring.

1.1. Statement of the Problem

Many scholars recommend the need to conduct in-depth studies on the risk factors of age at marriage among women for both developing and developed countries. Age at early marriage is a health issue as well as a human right violation. A recent review show that girls who marry before the age of 18 are disproportionately affected by complicated pregnancies that may lead to maternal mortality and morbidity: girls aged 10–14 are five times more likely to die in pregnancy or childbirth than women aged 20–24; girls aged 15–19 are twice as likely to die (UNICEF, 2011). A pregnancy too early in life before a girl's body is not fully mature is a major risk to both mother and baby. Also, they are more likely to experience complications of childbirth

including obstetric fistula and hemorrhaging (IWHP, 2009).

Mortality rates for babies born to mothers under age 20 are almost 75 percent higher than for children born to older mothers in Ethiopia. Teenage women are also twice as likely as older women to die due to complications during pregnancy and childbirth. Infants born from teenage mothers are more likely to suffer from low birth weight, and are at higher risk of dying in its first year by 60% compared with infants of mothers in their twenties (Nour, 2006). Age at first marriage has health implication for women and their under-five children (Adebawale, 2012).

In Ethiopia there is no studies that documented on the area of age at first marriage by using parametric frailty models except the studies were conducted on the early marriage by using logistic regression. Many of the studies used logistic regression analysis and Cox proportional hazard models to estimate the effect of covariates on the age at first marriage; which restricts attention to the events that occur within the shortest time observed and the correct inference based on Cox's models needs identically and independently distributed samples respectively. Logistic regression does not account the censoring observations i.e., does not hold for time-to-event data; however, survival analysis is more powerful than Logistic framework that takes censoring into considerations.

A frailty model is a generalization of a survival regression model and it accounts for the presence of an unobserved multiplicative effect on the hazard function by specifying independence among observed data items conditional on a set of unobserved or latent variables. However, the Cox proportional hazards model has no such constraint and the dependence of the event times is not accounted. The shared frailty model is used with multivariable survival data where unobserved frailty is shared within groups of individuals, and thus a shared frailty model may be thought of as a random effects model for survival data. However, different dependence structures result from different frailty distributions (Hougaard, 2000).

The study focuses on the modeling and identifying the impact of demographic and socio-economic factors on age at first marriage. The research questions are:-

- What are the key socio-economic and demographic predictors of age at first marriage amongst women in Ethiopia?
- Which baseline distributional assumption among the exponential, Weibull and log-logistic; as well as frailty distributions, the gamma and inverse Gaussian distributions well describe the age at first marriage?

1.2. Significance of the Study

The result of this study provides information on marriage in Ethiopian women by analyzing the impact of different variables on survival of age at first marriage.

Specifically;

- The results are expected to give some knowledge about the determinants or risk factors of age at first marriage in Ethiopian women.
- This study could be used as a landmark for further

studies related to marriage and others.

- This study could provide information to government and other concerned bodies in setting policies and strategies.

1.3. Objective of the Study

The specific objectives of this study are to:-

- Identify significant factors or covariates that are associated with time-to-age at first marriage for Ethiopian women.
- Determine parametric baseline hazard, which is appropriate in modeling the determinants of age at first marriage.

2. Materials and Methods

2.1. Data source

The data set in this study was obtained from Demographic and Health Survey data conducted in Ethiopia in 2011, which was the third comprehensive survey conducted as part of the worldwide Demographic and Health Surveys project. The data provide in-depth information on marriage, fertility, family planning, infant, child, adult and maternal mortality, maternal and child health, gender, nutrition, malaria, knowledge of HIV/AIDS and other sexually transmitted diseases.

2.2. Sample Design

The 2011 EDHS sample was selected using two stage cluster design and census enumeration areas (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. After a certain rearrangement, reorganization and removal of missing values the total number of women with complete information became 12,208.

2.3. Variables in the Study

2.3.1. The Response (Dependent) Variable

The dependent (outcome) 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. During the survey all women were asked a series of questions regarding to their marital status and whether they had ever lived with a man. The response to this question constitutes the women age at first marriage and women who had not yet experienced the events resulting in right censoring of the data.

2.3.2. Predictor (Independent) Variables

Several predictors are considered in this study to investigate the determinant factors for the timing of age at first marriage. All of these variables are categorical. Those

are respondents work status, religion, type of residence, head/parents education level, women education level, head/parents occupations, media exposure and wealth index.

Table 1. Operational definition and categorization of the covariate variables, EDHS, 2011.

Variables	Definition and Categorization
Women education	Women level of education (0= No education; 1= Primary; 2= Secondary and 3=Higher)
Residence	Place of residence for women (1=Rural; 2=Urban)
Wealth index	Household wealth index (1= Poor; 2=Medium; 3=Rich)
Religion	Women's religion (1= Orthodox; 2= Muslim; 3= Protestant; 4= Others)
Head/parents education	Education level of head/parents (0=No education; 1= Primary; 2= Secondary and 3=Higher)
Media	Access to media (0= No; 1= Yes)
Head/parents occupations	Occupational status of head/parents (1= Agriculturalists; 2= Professional; 3= Laborers; 4= Business; 5= Others)
Respondent work status	Working status of the respondent (0= Yes; 1= No)

Regional state of the women would be considered as a clustering effect in all frailty models.

2.4. Methodology: Survival Data Analysis

2.4.1. Non-parametric Methods

Suppose t_1, t_2, \dots, t_n be the survival times of n independent observations and $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(m)}$, $m \leq n$ be the m distinct ordered marriage times. The Kaplan-Meier estimator of the survivorship function (or survival probability) at time t , $S(t) = P(T \geq t)$ is defined as:

$$\hat{S}(t) = \prod_{t(j) \leq t} \left[\frac{n_j - r_j}{n_j} \right] = \prod_{t(j) \leq t} \left[1 - \frac{r_j}{n_j} \right]$$

With the convention that $\hat{S}(t) = 1$ for $t < t(1)$. In this equation, n_j is the number of individuals who are at risk of marriage at time t_j , and r_j is the number of individuals who occurs an event at time t_j .

The Cumulative hazard function of the KM estimator can be estimated as:

$$\hat{H}(t) = -\ln[\hat{S}(t)], \text{ where } \hat{S}(t) \text{ is KM estimator}$$

2.4.2. Shared Frailty Model

The frailty approach is a statistical modeling concept which aims to account for heterogeneity, caused by unmeasured covariates. In statistical terms, a frailty model is a random effect model for time-to-event data, where the random effect (the frailty) has a multiplicative effect on the baseline hazard function (Wienke *et al.*, 2003). Vaupel *et al.* (1979) used the frailty approach to derive the individual hazard function based on the population hazard function obtained from life tables. The shared frailty approach assumes that all failure times in a cluster are conditionally independent given the frailties. The value of the frailty term is constant over time and common to all individuals in the cluster, and thus it is responsible for creating dependence between event times in a cluster. This dependence is always positive in shared frailty models.

Conditional on the random effect, called the frailty denoted by u_i , the survival times in cluster i ($1 \leq i \leq n$) are assumed to be independent and the proportional hazard frailty model assumes:-

$$h_{ij}(t/x_{ij}, u_i) = h_0(t) \exp(\beta' x_{ij} + u_i)$$

where i indicates the i^{th} cluster and j indicates the j^{th} individual for the i^{th} cluster, $h_0(\cdot)$ is the baseline hazard function, u_i the random term of all the subjects in cluster i , X_{ij} the vector of covariates for subject j in cluster i , and β the vector of regression coefficients.

If the proportional hazards assumption does not hold, the accelerated failure time frailty model which assumes:-

$$h_{ij}(t/x_{ij}, u_i) = h_0 \exp(\beta' x_{ij} + u_i) \exp(\beta' x_{ij} + u_i)$$

If the number of subjects n_i is 1 for all groups, the univariate frailty model is obtained (Wienke, 2010); otherwise the model is called the shared frailty model (Hougaard, 2000; Duchateau and Janssen, 2008) because all subjects in the same cluster share the same frailty value.

Let us assume $Z = \exp(u_i)$ and assume Z has the gamma or the inverse Gaussian distribution, so that the hazard function depends upon this frailty that acts multiplicatively on it. Shared frailty models are very important in analyzing multivariate or clustered survival data. Shared frailty model assumes that all individuals in a subgroup or pair share the same frailty Z_i ($i=1, 2, \dots, n$), and because of this it is called shared frailty model, but frailty from group to group may differ. Shared frailty model is similar to the individual frailty model except the only difference is that frailty is now shared among the n_i observations in the i^{th} group.

$$h_{ij}(t) = Z_i h_0(t) \exp(\beta' X_{ij})$$

Baseline Survivor and Hazard Function

Let T be a random variable associated with the survival times, t be the realization of the random variable T and $f(t)$ be the underlying probability density function of the survival time t . The cumulative distribution function $F(t)$, which represents the probability that a subject selected at random will have a survival time less than some stated value t , is given by:

$$F(t) = P(T \leq t) = \int_0^t f(u) du, t \geq 0$$

The survivor function, denoted by $S(t)$, is defined to be the

probability of an individual surviving or being event-free beyond time t (experiencing the event after time t). It is defined as $S(t) = P(T > t) = \int_t^\infty f(u)du$. The survival function is merely the complement of the cumulative distribution function, that is $S(t) = 1 - F(t)$ and density function is:-

$$f(t) = \frac{-dS(t)}{dt} \quad t > 0$$

The hazard function is a measure of the probability of failure during a very small interval, assuming that the individual has survived at the beginning of the interval. It is defined as:-

$$h(t) = \frac{f(t)}{s(t)} = \frac{-d \ln s(t)}{dt}$$

Survival model is usually expressed in terms of hazard function. The cumulative hazard function is defined as:-

$$H(t) = \int_0^t h(u)du$$

Under the parametric approach, the baseline hazard is defined as a parametric function and the vector of its parameters are estimated together with the regression coefficients and the frailty parameter(s).

(i). Baseline Exponential Distribution

The exponential distribution, with only one unknown parameter and it is the simplest of all life distribution models. In the exponential model, the conditional probability is constant over time. In other words, the main feature of exponential distribution is that the instantaneous hazard does not vary over time. Modeling the dependency of the hazard rate on covariates entails constructing a model that ensures a non-negative hazard rate (or non-negative expected duration time). The exponential PH model is a special case of the Weibull model when $\gamma = 1$. The hazard function under this model is to assume that it is constant over time.

Table 2. Baseline Exponential distribution for survival and hazard functions, EDHS, 2011.

$f(t)$	$s(t)$	$h(t)$	$H(t)$	Parameter space
$\lambda \exp(-\lambda t)$	$\exp(-\lambda t)$	λ	λt	$\lambda > 0$

(ii). Baseline Weibull Distribution

Weibull distribution is one of the parametric distributions which are used for the analysis of life time data and mostly used in literature for modeling life time data (Ibrahim *et al.*, 2001 and Yu, 2006). The Weibull distribution is more general and flexible than the exponential distribution and allows for hazard rates that are non-constant but monotonic. It is a two-parameter model (λ and γ), where λ is the scale parameter and γ is the shape parameter because it determines whether the hazard is increasing, decreasing, or constant over time i.e., the hazard rate increases when $\gamma > 1$ and decreases when $\gamma < 1$ as time goes on. When $\gamma = 1$, the hazard rate remains constant, which is the special case of exponential.

Table 3. Baseline Weibull distribution for Survival and Hazard functions, EDHS, 2011.

$f(t)$	$s(t)$	$h(t)$	$H(t)$	Parameter space
$\gamma \lambda t^{\gamma-1} \exp(-\lambda t^\gamma)$	$\exp(-\lambda t^\gamma)$	$\gamma \lambda t^{\gamma-1}$	λt^γ	$\lambda, \gamma > 0$

(iii). Baseline Log-logistic Distribution

The cumulative distribution function can be written in closed form is particularly useful for analysis of survival data with censoring (Bennett, 1983). 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 is the scale parameter and γ is the shape parameter which is denoted by $\log L$ (γ, λ). The distribution imposes the following functional forms on the density, survival, hazard and cumulative hazard function:

Table 4. Baseline Log-logistic distribution for Survival and Hazard functions, EDHS, 2011.

$f(t)$	$s(t)$	$h(t)$	$H(t)$	Parameter space
$\frac{\lambda \gamma t^{\gamma-1}}{(1 + \lambda t^\gamma)^2}$	$\frac{1}{1 + \lambda t^\gamma}$	$\frac{\lambda \gamma t^{\gamma-1}}{1 + \lambda t^\gamma}$	$\ln \left[\frac{1}{1 + \left(\frac{t}{\lambda}\right)^\gamma} \right]$	$\lambda \in R, \gamma > 0$

By specifying one of the four functions $f(t)$, $S(t)$, $h(t)$ or $H(t)$ specifies the other three functions of the above baselines. The parameter λ is reparameterized in terms of predictor variables and the regression parameters. Typically for parametric models, the shape parameter γ is held fixed.

2.5. Frailty Distribution

2.5.1. Shared Gamma Frailty Distribution

The gamma distribution is very-well known and has simple densities. It is the most common distribution used for describing frailty. Even though gamma models have closed form expressions for survival and hazard functions, from a computational view, it fits well to frailty data and it is easy to derive the closed form expressions for unconditional survival and hazard functions.

To make the model identifiable, we restrict that expectation of the frailty equals one and variance be finite, so that only one parameter needs to be estimated. Thus, the distribution of frailty Z is the one parameter gamma distribution. Under the restriction, the corresponding density function and Laplace transformation of gamma distribution:-

$$f_Z(z_i) = \frac{z_i^{(1/\theta)-1} \exp(-z_i/\theta)}{\Gamma(1/\theta) \theta^{1/\theta}}, \theta > 0$$

where $\Gamma(\cdot)$ is gamma function. It corresponds to a Gamma distribution $\text{Gam}(\mu, \theta)$ with μ fixed to 1 for identifiability and its variance is θ .

The associated Laplace transform is:-

$$L(s) = (1 + s/\theta)^{-\theta}, \theta > 0$$

Note that if $\theta > 0$, there is heterogeneity. So the large values of θ reflect a greater degree of heterogeneity among

groups and a stronger association within groups. The conditional survival function of the gamma frailty distribution is given by: (Gutierrez, 2002).

$$S_{\theta}(t) = [(1 - \theta \ln\{S(t)\})]^{-1/\theta} \quad \theta > 0$$

The conditional hazard function of the gamma frailty distribution is given by: (Gutierrez, 2002)

$$h_{\theta}(t) = h(t)[1 - \theta \ln\{S(t)\}]^{-1}$$

where $S(t)$ and $h(t)$ are the survival and the hazard functions of the baseline distributions.

Larger variance indicates a stronger association within groups. For the Gamma distribution, the Kendall's Tau (Hougaard, 2000), which measures the association between any two event times from the same cluster in the multivariate case. It is an overall measure of dependence and independent of transformations on the time scale and the frailty model used. The associations within group members are measured by Kendall's, which is given by:-

$$\tau = \frac{\theta}{(\theta+2)}, \text{ where } \tau \in (0,1)$$

2.5.2. Inverse Gaussian Shared Frailty Distribution

Similar to the gamma frailty model, simple closed-form expressions exist for the unconditional survival and hazard functions, this makes the model attractive. The probability density function of an inverse Gaussian shared distributed random variable with parameter $\theta > 0$ is given by

$$f_Z(z_i) = \left(\frac{1}{2\pi\theta}\right)^{1/2} z_i^{-3/2} \exp\left(\frac{-(z_i-1)^2}{2\theta z_i}\right), \quad \theta > 0, z > 0$$

$$\begin{aligned} l_{\text{marg}}(h_o, \beta, \theta; Z, X) &= \prod_{i=1}^s \left[\left(\prod_{j=1}^{ni} (h_o(y_{ij}) \exp(x_{ij}^T \beta))^{\delta_{ij}} \right) X \int_0^{\infty} Z_i^{di} \exp\left(-Z_i \sum_{j=1}^{ni} H_o(y_{ij}) \exp(x_{ij}^T \beta)\right) f(z_i) dz_i \right] \\ &= \prod_{i=1}^s \left[\left(\prod_{j=1}^{ni} (h_o(y_{ij}) \exp(x_{ij}^T \beta))^{\delta_{ij}} \right) X (-1)^{di} L^{(di)} \left(\sum_{j=1}^{ni} H_o(y_{ij}) \exp(x_{ij}^T \beta) \right) \right] \end{aligned}$$

Taking the logarithm, the marginal likelihood is

$$l_{\text{marg}}(h_o(\cdot), \beta, \theta; Z, X) = \sum_i \{ [\sum_{j=1}^{ni} \delta_{ij} (\log(h_o(y_{ij})) + x_{ij}^T \beta)] + \log[(-1)^{di} L^{(di)}(\sum_{j=1}^{ni} H_o(y_{ij}) \exp(x_{ij}^T \beta))] \}$$

where $di = \sum_{j=1}^{ni} \delta_{ij}$ is the number of events in the i^{th} clusters and $L^{(q)}(\cdot)$ is the q^{th} derivative, the Laplace transform of the frailty distribution Z is defined as:-

$$\begin{aligned} L(s) &= E[\exp(-Zs)] = \int_0^{\infty} \exp(-Zs) f(Z_i) dz_i, s > 0 \text{ and} \\ L^{(q)}(s) &= (-1)^q \int_0^{\infty} Z^q \exp(-Zs) f(z) dz, q \geq 0 \end{aligned}$$

where $h_o(\cdot)$ represents a vector of parameters of the baseline hazard function, β the vector of regression coefficients and θ the variance of the random effect. The estimates of $h_o(\cdot)$, β ,

For identifiability, we assume z has expected value equal to one and variance θ .

The Laplace transformation of the inverse Gaussian distribution is:-

$$L(s) = \exp\left[\frac{1-(1+2\theta s)^{1/2}}{\theta}\right] \quad \theta > 0, s > 0$$

For the inverse Gaussian frailty distribution the conditional survival function is given by: (Gutierrez, 2002).

$$S_{\theta}(t) = \exp\left\{\frac{1}{\theta} \left(1 - [1 - 2\theta \ln\{S(t)\}]^{1/2}\right)\right\} \quad \theta > 0$$

For the inverse Gaussian frailty distribution the conditional hazard function is given by: (Gutierrez, 2002).

$$h_{\theta}(t) = h(t)[1 - 2\theta \ln\{S(t)\}]^{-1/2} \quad \theta > 0$$

where $S(t)$ and $h(t)$ are the survival and the hazard functions of the baseline distributions.

With multivariate data, an Inverse Gaussian distributed frailty yields a Kendall's Tau given by:-

$$\tau = \frac{1}{2} - \frac{1}{\theta} + 2 \frac{\exp(2/\theta)}{\theta^2} \int_{2/\theta}^{\infty} \frac{\exp(-u)}{u} du, \text{ where } \tau \in (0, 1/2)$$

2.6. Method of Parameter Estimation

Under the assumption of right-censoring and of independence between the censoring time and the survival time of random variables, given the covariate information, the marginal log-likelihood of the observed data can be

θ are obtained by maximizing the marginal log-likelihood of the above. This can be done if one is able to compute higher order derivatives $L^{(q)}(\cdot)$ of the Laplace transform up to $q = \max\{d_1, \dots, d_s\}$. Symbolic differentiation is performed in R, but is impractical here; mainly because this is very time consuming Munda *et al.* (2012).

2.7. Comparison of Models

Model comparison and selection are among the most common problems of statistical practice, with numerous procedures for choosing among a set of models (Kadane and Lazar, 2001) and (Rao and Wu, 2001). There are several

methods of model selection. The most commonly used methods include information criteria. One of the most commonly used model selection criteria is Akaike Information Criterion (AIC). A data-driven model selection method such as an adapted version of Akaike's information criterion AIC (Akaike, 1974) is used to find the truncation point of the series. In some circumstances, it might be useful to easily obtain AIC value for a series of candidate models (Munda *et al.*, 2012). In this study, we used the AIC criteria to compare various candidates of parametric frailty models. The model with the smallest AIC value is considered a better fit.

2.8. Model Diagnostics (Checking)

2.8.1. Evaluation of the Baseline Parameters

The graphical methods can be used to check if a parametric distribution fits the observed data or not. The model with the Weibull baseline has a property that the $\log(-\log(\hat{S}(t)))$ is linear with the $\log(t)$, where $\hat{S}(t) = \exp(-\lambda t^\gamma)$. Hence, $\log(-\log(\hat{S}(t))) = \log(\lambda) + \gamma \log(t)$. The intercept and slope of the line will be rough estimate of $\log \lambda$ and γ respectively. 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 Timon Stucki, 2011).

The appropriateness of the model with the exponential baseline can graphically be evaluated by plotting $-\log(\hat{S}(t))$ versus t where $\hat{S}(t)$ is Kaplan-Meier survival estimate. This plot should be linear and goes through the origin (Klein, 1992). Because for exponential distribution, $\hat{S}(t) = \exp(-\lambda t)$, and hence, $-\log(\hat{S}(t)) = \lambda t$ is linear with time.

The appropriateness of the model with the log logistic baseline can graphically be evaluated by plotting $\log(1 - \hat{S}(t)/\hat{S}(t))$ versus $\log(t)$, where $\hat{S}(t)$ is Kaplan-Meier survival estimate. The log-failure odd versus log time of the log-logistic model is linear with slope γ then the survival time follows a log-logistic distribution. Where the failure odds of log-logistic survival model can be computed as:

$$1 - s(t)/s(t) = \frac{\lambda t^\gamma}{1 + \lambda t^\gamma} = \lambda t^\gamma$$

Therefore, the log-failure odds can be written as

$$\log((1 - s(t))/s(t)) = \log(\lambda t^\gamma) = \log(\lambda) + \gamma$$

which is the liner function of $\log(t)$. (Dätwyler and Timon Stucki, 2011).

2.8.2. The Cox-Snell Residuals

The Cox-Snell residuals method can be applied to any parametric model and the residual plots can be used to check the goodness of fit of the model. 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 and Moeschberger, 2003).

The Cox-Snell residual for the j^{th} individual with observed survival time t_j is given by $r_j = \hat{H}(T_j/X_j) - \log \hat{S}(T_j/X_j)$, where \hat{H} and \hat{S} are the estimated values of the cumulative hazard and survivor function of the j^{th} subject at time t_j respectively. If the model fits the data, then the r_j 's should have a standard ($\lambda=1$) 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 unity and zero intercept. If yes, the fitted model is adequate. In general, Cox-Snell residual that provides a check of the overall fits of the model (Cox and Snell, 1968).

The three baseline hazard functions of Cox-Snell residuals that are considered in this study are given below:

Table 5. The three baseline hazard functions of Cox-Snell residuals, EDHS, 2011.

Model	r_j
Exponential	$\lambda t_j \exp(\beta' X_j)$
Weibull	$\lambda t_j^\gamma \exp(\beta' X_j)$
Log-logistic	$\frac{1}{1 + \lambda t_j^\gamma \exp(\beta' X_j)}$

3. Results

3.1. Summary Statistics

A total of 12208 women's was included in the study during the data collection. From the total 8462(69.3%) were experienced in marriage and the rest of 3746(30.7%) were unmarried at different age of marriage between 8 years and 49 years. Furthermore, among 60.6% of women were married before age of 18 years. This indicates that age at early marriage is highest in Ethiopia. Regarding to educational attainment, about 48.4% of heads/parents were illiterate while 34.0% of the heads/parents had attended primary education and the remaining 17.6% was secondary and higher education. About 60.4% of the women respondents had no work and 39.6% had a work. Women who were residing in rural area were 67.6% whereas women who residing in urban area were 32.4%. About 46.6% of the household's wealth indexes were classified as poor while 18.5% had medium income and 34.9% were rich.

The study revealed that educational attainments of women; about 51.3% had no education while 36.2% had primary education and the remaining 12.5% had attended secondary and higher education. From the total number, 63.1% of heads/parents occupations were agriculturalist, 17.6% were professional, 12.4% were Business and 6.9% were Laborers and Others. With regard to exposure to mass media, 50.6% of the women respondent had no any access of media and 49.4% was recorded for women who had any access of mass media. The minimum and maximum ages at first marriage in the data were 8 years and 37 years respectively. The median of age at first marriage was 17 years. The skewness of age at first marriage was 1.372. This

shows a data is skewed to the right distribution with the 25th, 50th and 75th percentiles of age at first marriage was 15, 17 and 19 years respectively.

The log-logistic-inverse Gaussian shared frailty model that is 50131.9497 appears to be appropriate model compared with other models. This indicates it is more efficient model to describe age at first marriage dataset.

Table 6. The value of AIC for Multivariable Parametric Shared Frailty Models, EDHS, 2011.

Model		AIC
Baseline Hazard function	Frailty distribution	
Exponential	Gamma	71260.6561
	Inverse Gaussian	71260.7655
Weibull	Gamma	52496.8809
	Inverse Gaussian	52496.8240
Log-logistic	Gamma	50133.8556
	Inverse Gaussian	50131.9497

3.2. Multivariable Analysis

The variance of the frailty were significant for all baseline hazard function with an inverse Gaussian shared frailty distribution in the models whereas it was not significant in the gamma shared frailty distribution using the same baseline as inverse Gaussian models at 5% level of significance. This indicates the presence of heterogeneity and necessitates the frailty models. The value of shared frailty distribution (θ) is 0.029, 0.416 and 0.312 for exponential-inverse Gaussian, Weibull-inverse Gaussian and log-logistic-inverse Gaussian

respectively. In this case it is highest when Weibull was used as baseline next to log-logistic and it is smaller when exponential was used as baseline. The dependence within clusters (region) for the exponential-inverse Gaussian shared frailty model ($\tau=0.014$) and Weibull-inverse Gaussian shared frailty model ($\tau=0.134$).

Model comparisons were presented in Table 5. Accordingly, it suggested that log-logistic-inverse Gaussian shared frailty model was selected according to AIC. In this model all categorical variables were significant except some category of religion and wealth index of the household. From Table 6 the confidence intervals of the acceleration factor for all significant categorical covariates do not include one at 5% level of significance. This shows that they were significant factors for determining the time-to-age at first marriage for Ethiopian women. However, from the variable of religion category Muslims were not significant when using orthodox as the reference category with (p-value=0.052, $\phi = 1.0113$, 95%CI = (0.9998, 1.0228) and chi-sq = 3.78). Also, those households who had middle wealth index was insignificant by using poor households as a reference category with (p-value = 0.74, chi-sq= 0.11, 95%CI = (0.9911, 1.0127) and $\phi = 1.0018$). The estimated coefficient of the parameters for respondent who had no work status was -0.0116. The sign of the coefficients are negative which implies that decreasing logged of survival time and hence, shorter expected duration of age at first marriage of the women.

Table 7. Results of the multivariable Log-logistic-Inverse Gaussian shared Frailty Model for age at first marriage dataset, EDHS, 2011.

Variable	Coef	S.e(coef)	ϕ	95%CI LCL UCL	Chi-sq	P-value
(Intercept)	2.7976	0.0324	16.4052	(15.3958,17.4808)	7459.52	0.00
Heads/parents occupation						
Agriculturalist	Ref					
Professional	0.0216	0.0070	1.0218	(1.0079,1.0360)	9.45	2.1e-13
Business	0.0433	0.0067	1.0433	(1.0306,1.0581)	42.17	8.3e-11
Laborers	0.0432	0.0115	1.0441	(1.0209,1.0679)	14.04	1.8e-04
Others	0.0524	0.0118	1.0538	(1.0297,1.0785)	19.91	8.1e-06
Women Education						
No education	Ref					
Primary	0.0118	0.0047	1.0119	(1.0026,1.0212)	6.26	0.012
Secondary	0.0766	0.0089	1.0796	(1.0609,1.0986)	73.66	0.00
Higher	0.1395	0.0112	1.1497	(1.1247,1.1752)	155.9	0.00
Residence						
Rural	Ref					
Urban	0.0169	0.0067	1.0170	(1.0038,1.0305)	6.45	0.011
Heads/parents Education						
No education	Ref					
Primary	0.0344	0.0080	1.0350	(1.0189,1.0514)	18.51	1.7e-05
Secondary	0.0397	0.0047	1.0405	(1.0310,1.0501)	69.94	1.1e-16
Higher	0.0434	0.0098	1.0444	(1.0245,1.0646)	19.51	9.9e-06
Media						
No	Ref					
Yes	0.0200	0.0046	1.0202	(1.0110,1.0294)	18.67	1.6e-05
Religion						
Orthodox	Ref					
Muslims	0.0112	0.0058	1.0113	(0.9998,1.0228)	3.78	0.052
Protestant	0.0213	0.0071	1.0215	(1.0074,1.0358)	8.79	0.0027
Others	0.0309	0.0138	1.0314	(1.0039,1.0597)	5.03	0.025
Wealth Index						
Poor	Ref					
Middle	0.0018	0.0055	1.0018	(0.9911,1.0127)	0.11	0.74

Variable	Coef	S.e(coef)	ϕ	95%CI LCL UCL	Chi-sq	P-value
Rich	0.0477	0.0053	1.0489	(1.0380,1.0598)	82.85	0.00
Work status						
Yes	Ref					
No	-0.0116	0.0043	0.9885	(0.9802,0.9968)	7.25	0.0071
Log(scale)	-2.1305	0.0090	0.00			
Frailty	104.42	0.00				
	$\theta = 0.312$	$\lambda = 6.1669\text{e-}11$				
	$\tau = 0.109$	$\gamma = 8.4034$	AIC= 50131.9497			

Coef= coefficient, S.e= standard error, ϕ = acceleration factor, 95% CI=Confidence Interval for acceleration factor, LCL=lower class limit, UCL= upper class limit, Chi-sq= Chi-square, Ref=Reference, θ = variance of the random effect, λ = scale parameter, γ = shape parameter, τ = Kendall's Tau.

The occupational status of heads/parents was statistically determine age at first marriage of the women. The time rate and 95% Confidence interval of acceleration factors for occupational status of heads/parents for a group of professional, business, laborers and Others was 1.0218(1.0079, 1.0360), 1.0433(1.0306, 1.0581), 1.0441(1.0209, 1.0679) and 1.0538 (1.0297, 1.0785) when compared to occupation of agriculturalists group (as reference category) respectively.

The 95% confidence interval for acceleration factor of women educational levels was (1.0026, 1.0212), (1.0609, 1.0986) and (1.1247, 1.1752) for the group of primary, secondary and higher education's respectively. This confidence interval does not include one in all; 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. Accordingly, it prolonged the age at first marriage by a factor of ($\phi = 1.0119$, $\phi = 1.0796$ and $\phi = 1.1497$) for primary, secondary and higher education respectively at 5% level of significance.

The acceleration factor for women who are lived in urban area was 1.0170 times greater than those who are lived in rural area (Ref) (ϕ : 1.0170, 95%CI: 1.0038, 1.0305). The 95% confidence interval of the acceleration factor for those women who had any access of media and don't had works are (1.0110, 1.0294) and (0.9802, 0.9968), its acceleration factors are 1.0202 and 0.9885 by using those not had any access of media and had a work as a reference category at 5% level of significance respectively.

The coefficients on the categorical variable of heads/parents educational level, shows that the survival time of age at first marriage increased with changing from one category to another (primary, secondary and higher) educational level relative to those heads/parents with no education as a reference group and the survival times was lengthened by 3.50%, 4.05% and 4.44% respectively for the group of primary, secondary and higher educational level of heads/parents.

The estimate of shape parameter in the log-logistic-inverse Gaussian shared frailty model is ($\gamma=8.4034$). This value shows the shape of hazard function is unimodal because the value is greater than unity i.e., it increases up to some time and then decreases. The heterogeneity in the population of the region which is used as a clusters are estimated by our selected model is $\theta=0.312$ and the dependence within the

clusters (region) is measured by Kendall's tau is $\tau=0.109$.

3.3. Survival Function of Different Categorical Group of Covariates

In all frailty models the categories of women's educational level were highly significant at 5% level of significance when compared with the reference category (no education). The gap between the four curves distinguishes that the survival distribution of age at first marriage for Ethiopian women. The differences that are displayed in survival curve emphasize that women who had higher education was married later when compared with others and women who are uneducated had less survival than educated. As we indicate from the graph there is a high gap at the mid time between a primary and secondary educated woman on marriage. This shows women who had secondary education were more survived than uneducated and primary educated women.

Survival Function womens education level vs Time(years)

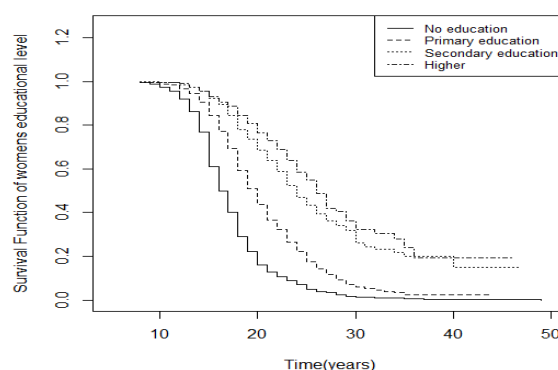


Figure 1. Displays the survival function for the group of women educational level on age at first marriage by using log-logistic-inverse Gaussian shared frailty model, EDHS, 2011.

3.4. Discussion of Results

The findings of this study revealed that the educational level of women had a significant effect on the survival of age at first marriage with 5% level of significance and it prolonged age at first marriage by the factor of $\phi= 1.0119$, 1.0796 and 1.1497 for primary, secondary and higher education respectively when illiterate women was used as the reference group. The result of the study shows that woman who had higher education was more survived than those

uneducated and primary education. A study conducted in Ethiopian regions by Erulkar (2011) investigated the factors associated with marriage and the result suggested that educational attainments of women had significant effect on marriage and women who were not educated were married earlier than educated.

The results of this study suggested that place of residences was significant predictive factor for age at first marriage in Ethiopian women. This shows that women who lived in urban areas are more survived on age at first marriage than women who lived in rural areas. This might be rural areas tend to have institutional and normative structures such as the kinship and extended family that promote early marriage and childbearing, but 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. A study in Nigeria by Thomas (2010-2011) and Adebawale (2012) found that women who are lived in rural area had a higher risk of first marriage than urban area i.e., hazard of women living in rural is greater than urban.

The results of this study suggest that work status of the women had a significant effect on age at first marriage and age at early marriage was higher for women who had no work status. This is consistent with Shapiro (1996), Zahangir *et al.* (2008) and Kamal (2011), they revealed that work status of women were significant effect on age at first marriage and pre-marital work status of women were significantly delayed the timing of marriage.

Households who had higher wealth index were found to be one of the statistically significant factors from the category for determining age at first marriage in our study. It showed that age at first marriage for women was prolonged by the factor of $\phi = 1.0489$ when we used the wealth index of poor household as (reference). This finding is consistent with Kamal (2011) in Tribal women in Bangladesh and the study revealed that the higher economic status of parents, the lower is the probability of age at early marriage.

The result of this study also revealed that heads/parents occupations are the important factor for age at first marriage of Ethiopian women. A similar study in Western Uganda by Peninah *et al.* (2011) and Zahangir *et al.* (2008) in rural Bangladesh found that the occupation of the parents were strong socio-economic determinants of age at first marriage. Also, another study in Bangladesh by Mosammat *et al.* (2013) showed that the occupations of the parents were important factors for determining age at first marriage.

Access to mass media was found to have a significant effect on age at first marriage. The findings of this study showed that women who had no access of media were married at earlier age than those who had access of media. This finding is consistent with Tezera (2013), Zahangir and Kamal (2011), Zahangir *et al.* (2008) and Joseph *et al.* (2012).

In this study, we used the region as a clustering (frailty) effect on modeling the determinants of time-to-age at first marriage in Ethiopian women using 2011 EDHS data. The clustering effect were significant (p -value < 0.000) in log-logistic-inverse Gaussian shared frailty model. This showed

that there is heterogeneity between regions by assuming women within the same region share similar risk factors on marriage i.e., the correlation within regions cannot be ignored and clustering effect was important in modeling the hazard function.

In our study the adequacy of baseline distributions are checked by using graphs in figure (5). From the plot of exponential, Weibull and log-logistic distributions; the plot of log-logistic was more straight line compared with exponential and Weibull for age at first marriage dataset. These findings were consistent with Cox (1970), O'Quigley and Struthers (1982), Bennet (1983) and Cox and Oakes (1984) for baseline log-logistic.

4. Conclusions

This study was based on a dataset of age at first marriage obtained from the central statistical agency of Ethiopia with an aim of modeling the determinants of time-to-age at first marriage by using different parametric baseline with different shared frailty model on the marriage dataset of Ethiopian women. Out of the total 12208, about 69.3% were experienced an event (married) and 30.7% were not experienced an event (unmarried) for a different age between 8 and 49 years.

To model the determinants of time-to-age at first marriage, various parametric shared frailty models by using different baseline distributions were applied. Among this using AIC, the log-logistic-inverse Gaussian shared frailty model is better fitted to marriage dataset than other parametric shared frailty models. There is a frailty (clustering) effect on the time-to-age at first marriage dataset that arises due to differences in distribution of time to age at first marriage among region of Ethiopia.

The result of log-logistic-inverse Gaussian shared frailty model showed that the factors that determine the timing of age at first marriage are women educational level, heads/parents occupation, place of residence, educational level of heads/parents, access to any media and work status of the respondents are statistically significant. Also from the category of religion, protestant and others and from household's wealth index, the richest households are significant. As educational level of the women increases, age at first marriage is highly prolonged in Ethiopian women. This indicates education of women were significant factor to determine timing of age at first marriage and implies that girls should be kept in school for a longer period, not only for the purpose of raising age at marriage, but also for biological, physical and mental maturity.

Awareness has to be given for the society on age at the marriage. The mass media can play an effective role in this regard and the awareness need to follow the ordinance of the legal age of marriage because it is the most determinants of health for women and child borne. Similarly, it is advisable to target young women, particularly those with no or little education including primary school girls, with information on reproductive health and to provide them to avoid ultimately

age early marriage. 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 government should takes an action on age at first marriage.

Acknowledgement

The authors of this article would like to the thank Central Statistical Agency of Ethiopia for making available the data used in this research. We also thank university of Gondar for the use of their computers and Internate services.

Reference

- [1] 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*.
- [2] Akaike H. (1974). A new look at the statistical model identification. *IEEE Trans Automatic Control*.
- [3] Bayisenge J. (2010). Early Marriage as a Barrier to Girl's Education Rwanda: National University of Rwanda.
- [4] Bennett S. (1983). Analysis of survival data by the proportional odds model. *Statistics in Medicine*, 2, 273.
- [5] Bedassa B. (2015). Risky Sexual Behavior and Predisposing Factors to HIV/STI Among Students in Mizan-Tepi University(Acase of Tepi Campus). *Science Journal of Public Health*, Vol.3, No.5, Pp.605-611.
- [6] Cox D. (1972). Regression models and life tables (with discussions). *Journal of the Royal Statistical Society*. 34: 187-220.
- [7] Cox D. R. and Oakes D. (1984). Analysis of Survival Data. Chapman and Hall, London.
- [8] Cox D. R. and Snell E. J. (1968). A general definition of residuals with discussion. *Journal of the Royal Statistical Society. Series B* 30 (1968), 248-275.
- [9] Duchateau L. and Janssen P. (2008). The Frailty Model. *Springer-Verlag*, New York.
- [10] Erulkar A. (2013). Early Marriage, Marital Relations and Intimate partner violence in Ethiopia.
- [11] Goncalves L., Duarte H. and Cabral M. (2015). Prevalence of Hemoglobin S in Blood Donors in the Hospital Dr. Agostinho Neto, Praia City-Cape Verde. *Science Journal of Public Health*, 3:5, 600-604.
- [12] Hougaard P. (1984). Life table methods for heterogeneous populations. *Biometrika* 71, 75-83.
- [13] Hougaard P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 387 - 396
- [14] Hougaard P. (1995). Frailty Models for Survival Data. *Lifetime Data Analysis*. 1: 255-273.
- [15] Hougaard P. (2000). Analysis of Multivariate Survival Data, *Springer-Verlag*, Newyork.
- [16] Ibrahim J.G., Chen M. and Sinha D. (2001). Bayesian survival analysis. Springer Verlag, New York.
- [17] IPPF and UNFPA. (2006). Ending Child Marriage: A Guide for Global Policy Action, International Planned Parenthood Federation (IPPF), London
- [18] 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.
- [19] Kadane J. and Lazar N. (2001). Methods and Criteria for Model Selection. Technical Report, 759, Carnegie Mellon University.
- [20] Kamal S.M. Mostafa (2011). Socio-Economic Determinants of Age at First Marriage of the Ethnic Tribal Women in Bangladesh, *Asian Population studies*.
- [21] Klein J. (1992). Survival analysis: techniques for censored and truncated data. *Medical College of Wisconsin*.
- [22] Klein P. and Moeschberger L. (2003). Survival Analysis Techniques for Censored and Truncated Data - Google Books.htm
- [23] Mosammat Zamilun Nahar, Mohammad Salim Zahangir and S.M. Shafiqul Islam (2013). Age at first marriage and its relation to fertility in Bangladesh, *Chinese Journal of Population Resources and Environment*, 11:3, 227-235, DOI:10.1080/10042857.2013.835539
- [24] Munda M., Rotolo F. and Legrand C. (2012). Parametric Frailty Models in R. *Journal of American Statistical Association*: 55:1-21.
- [25] Nour N. M. (2006). Health Consequences of Child Marriage in Africa. *Emergence Infection Disease* 12(11):1644-9.
- [26] Peninah A., Leonard K. Atuhaire. and Gideon Rutaremwa (2011). Determinants of age at first marriage among women in western Uganda. Presented in European Population Conference 2010, Vienna 1-4 September 2010
- [27] Population Council (2004). Supporting Married Girls: Calling Attention to a Neglected Group.
- [28] Rao C. and Wu Y. (2001). On Model Selection (with discussion). In P. Lahiri (Ed.), Model Selection, Volume 38 of IMS Lecture Notes - Monograph Series. Institute of Mathematical Statistics.
- [29] Shapiro, D. (1996). Fertility Decline in Kinshasa, *Population Studies* Vol50, No.1, pp 89- 103.
- [30] Tezera A. (2013). Determinants of Early Marriage among Women in Ethiopia.
- [31] Thomas E. (2010-2011). Multilevel Survival Analysis of the Determinants of Age at First marriage Among Women living in Nigeria. Thesis for Master of Statistical Data Analysis, Gent University.
- [32] UN . (2000). World Urbanisation Prospects. New York, United Nations.
- [33] UN. (1990). Patterns of First Marriage: Timing and Prevalence, *New yorke* Vol.4. No. 3 (April), pp: 221-235.
- [34] UN. (2007). The Millennium Development Goals, Report 2007, United Nations, New York. Retrieved 12, August, 2013 <http://www.un.org/millenniumgoals/pdf/mdg2007.pdf>

- [35] UNFPA. (2003). State of World Population 2003: Making 1 Billion Count: *Investing in Adolescents' Health and Rights*. New York
- [36] 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.
- [37] UNICEF. (2005). Early Marriage: A Harmful Traditional Practice: A Statistical Exploration. UNICEF: New York, NY.
- [38] UNICEF. (2011). Child Protection from Violence, Exploitation and Abuse.
- [39] Vaupel J., Manton K. and Stallard E. (1979). The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography*. 16: 439-454.
- [40] Westoff C. F. (2003). Trends in Marriage and Early Childbearing in Developing Countries. DHS Comparative Reports No.5. Macro International Inc.: Calverton, Maryland.
- [41] Wienke A., Lichtenstein P. and Yashin A.I. (2003). A bivariate frailty model with a cure fraction for modeling familial correlations in diseases. *Biometrics*, 59, 1178-83.
- [42] Wienke A., Ripatti S., Palmgren J., Yashin A.I. (2010). A bivariate survival model with compound Poisson frailty. *Statistics in Medicine* 29, 275-83.
- [43] Yu B. (2006). Estimation of shared gamma frailty models by a modified EM algorithm. *Computational Statistics and Data Analysis* 50, 463-474.
- [44] Zahangir M. S., M. A. Karim, M. R. Zaman, M. I. Hussain and M. S. Hossain, 2008. Determinants of age at first marriage of rural women in Bangladesh: A cohort analysis *Trends Applied Science. Res.* 4(3):335-343, *Academic journal, Department of Statistics*.
- [45] Zahangir M.S. and Kamal M.M. (2011). Several Attributes Linked with Child Marriage of Females' in Bangladesh, *International Journal of Statistics and Systems*, volume 6.