A Semi-Parametric Regression Hazards Model for Duration of Singlehood in North-East India

Lourembam Neroka Devi and Kshetrimayum Anand Singh
Department of Statistics
Manipur University,Canchipur, Imphal, Manipur, 795003, India

Abstract

Marriage is a life event which could change the qualitative status of an individual from single to married life. Studies on the duration of an individual's single status are of particular importance as it reflects the age pattern of marriage for a particular community or population. The median age at first marriage of India count to be 18.6 years for women and 24.5 years for men. This study aims the differentials and determinants of male and female singlehood durtation in North East States of India. Data from NFHS-4 are used to compute median duration of singlehood and its influential covariate are determined by using semi-parametric hazards model. Results show that the median duration of singlehood for North-East women and men are 21 years and 26 years respectively. Manipur women and men are recorded highest singlehood duration of 23 years and 27 years respectively. Findings shows that covariates such as place of residence, religion, ethnicity, wealth of the family and working status of women and men have significant effects on the duration of singlehood. As early marriages are expected to contribute more births it is important to increase the age at marriage of both men and women in order to reduce fertility.

Key words: Singlehood; NFHS-4; Median duration; Semi-parametric hazards model; North East India; Manipur.

Mathematics Subject Classification: 62G08, 91D20

1.   Background

Singlehood is a term used for an individual who has never been married in his or her lifetime. Studies on singlehood for both men and women are equally important in a society. All the life course activities of an individual during his/her singlehood are determining factors for the individual's future shape. The status of single for an individual is defined in terms of their relationship to marriage. Marriage is a major life event where a change of status takes place in an individual's life course whoever male or female. In many societies, marriage is defined as the onset of socially accepted sexual activity and as such marriage is considered as an important proximate determinant of fertility by Bongaarts (1978). And also it is the onset

Corresponding Author:Lourembam Neroka Devi
Email: lourembamneroka@gmail.com

of making one family and society sooner or later. Marriage is a demographic event which could change the nomenclature of one's social recognition or status as an individual that is from single status (bachelor or spinster) to married life. There is differences or changes in the life course activities and routines between a never married individual and an ever married individual.

Stein (1975) suggest that U.S. census and surveys indicate the increased postponement of marriage led the growing number of singles. Singlehood as a positive choice have been made by adults who have chosen not to marry. Due to dissatisfaction with traditional marriage, a new lifestyle of being single throughout their life has been chosen. In addition, Stein (1975) reveals that more and more people are rejecting and postponing their marriage in favour of independence.

Many women in developing countries of the world are subject to early marriage. It is believed that many women in such countries have little to no chance to choose themselves to whom they should marry and at what age they would marry by Jensen and Thornton (2003). Hayase and Liaw (1997) claim that women who marry at an early age have a longer period of exposure to pregnancy and consequently led to high fertility level. Jensen and Thornton (2003) also reveals that early married women face many disadvantages in the field of education, status, autonomy and even including physical safety. They have less power on decision-making, and better experiences of domestic violence are reported from them.

Kumchulesi *et al.* (2011) suggest that many socio economic factors such as age of women, place of residence, region, economic status *etc.* have an effect on the age at first marriage. With rapid increase in the educational attainment, age at first marriage and age at first birth is also increased. In addition, Gangadharan and Maitra (2000) also found that education of husband significantly affects the time to conception.

The study from Weinberger (1987) found that early marriage occurs more often in the less educated women. Findings from the world fertility survey 1987 which include 38 countries around the world, shows that singulate mean age at marriage (SMAM) of women with seven or more years of education is almost four years higher than the women with no education. And Matlabi *et al.* (2013) also suggest that one of the method for reducing early child marriage is mandating girls stay in school. Early marriage is associated with early childbearing and also linked to various adverse health related outcomes for both mother and child. Such early child bearing is lowered by increasing longer duration of singlehood with subsequently slowing population growth.

Kumar (2016) studies provide that place of residence is a responsible variable for a wide variation in child marriage. The percentage of girls married before 18 years of age among all those got married 0-4 years before to census 2011 of India in rural areas was 21% while it is only half in urban area. Other studies from McLaughlin *et al.* (1993); Westoff and ORC Macro (2003) also show that rates of early marriage are higher in rural areas than in urban areas.

A longer duration of singlehood results in lowering childbearing experience with subsequently slowing population growth. Women who have higher level of intelligence, education and occupation are more likely to remain as a single for a longer period of time. Highly educated women want to live on their own way. However, studies from Spreitzer and Riley

(1974), in contrast to women, educated men, and those with higher occupational achievement want to get married sooner.

Lalmalsawmzauva *et al.* (2011) claim that, though the legal age at marriage for girls in India is fixed at 18 years many girls are married before reaching that particular age. The female age at marriage in India is not uniform in all states, districts, ethnic, caste, class and religious groups. The female of rural areas get married earlier than those of urban. Some states located to southern part (except Andhra Pradesh), North West and North East India have relatively higher mean age at marriage .

Although, till date marriage is universal in Indian context, there are certain shifts observed in the age at marriage. There is a consistence increasing trend in respect of mean and median age at marriage over cohort since 2005 to 2016 (NFHS-3 (2007) and NFHS-4 (2017)). Thus it becomes important to understand the current situation of marriage pattern in India in the light of policies aimed at increasing the age at marriage and the major contribution factors determining the change in median age at marriage in the last decade. Not much work had been done to model duration of singlehood exclusively for North East India. Thus, the present study attempt to analyze the differentials and determinants of singlehood duration of North Eastern States of India.

2.   Data and methods

In view of literature reported on the age at marriage, the authors have an interest on the singlehood duration both for men and women the for whole North East India. The North Eastern region of India comprises of eight states *viz.* Assam, Arunachal, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura. These states have socio-economic and demographic characteristics, which are more or less different from the mainland population of India. Specifically, the economic activity in the region is quite different from the mainland as having little to no industrialization and mainly depends on agricultural activities. All states are dominated by tribal population except Assam where tribal population accounts for 12.5% only. The main religious groups in the region are Hindu, Muslim, Christians, Buddhists and some unrecognized local faiths still exist though fewer in number. The population of the region is sparse as compared to other parts of country and shares only 3.57% of whole population of India while the geographical area covers 7.5% of the total area of the country. All states except Meghalaya follow patrilineal norms while in Meghalaya there are ethnic groups who follow matrilineal norms. The present paper uses data from the National Family Health Survey-4 which provides information on various aspects of demographic analysis, reproductive health and nutrition for India. NFHS-4 (2017), 2015-16 (International Institute for Population Sciences (IIPS) and ICF, 2017) collected information from nationally representative samples regarding women, men, household and children. Interview was taken from 98702 women in the age group 15-49 years of the eight states of North Eastern region of India. Also, 14555 men in the age group 15-54 years were interviewed during the survey in the region. The duration of singlehood for these women and men were obtained from their age at marriage and current age. For those women or men who are never married, their duration of singlehood is obtained from their current age and is marked censored. Censored duration indicates that the event of interest (*i.e* marriage) does not occur to these women and men, whereas a complete duration indicates that the event has occurred at least once to the individuals. Out of total the women samples from North East India, 93321 women were

considered in the present study. Likewise, 14280 men in age group 15-54 were considered in the analysis.

2.1.   Variables in the study

In any regression analysis one has to ascertain the outcome and the predictor variables. The outcome or dependent variable in the present study is the duration of singlehood. In the event history method duration of singlehood may be looked upon as the time to first marriage. In the literature, time to first marriage is the duration of the total time where an individual lives in the single state starting from the birth of individual. Several predictor variables are considered in the present study which are potential to influence the duration of singlehood. All the variables are categorical variables. These variables which are thought to influence the outcome variable are educational level of individuals, type of place of residence, religion, ethnicity, wealth of family, exposure to mass media and working status. At the community level, place of residence, religion and ethnicity are considered. At the household level wealth of family is considered and at the individual level educational level, exposure to mass media and working status are included as covariates. Table 1 gives the definition and categories of the predictor variables.

2.2.   Methodology

The duration of singlehood (or time to first marriage) has been studied by way of survival analysis techniques using the non-parametric Kaplan-Meier, see Kaplan and Meier (1958) method and the semi-parametric Cox proportional hazards model, Cox (1972). As noted earlier duration of singlehood is the time an individual has got married for the first time starting from birth. In the present study, the event of interest is the first marriage. The Cox proportional hazards model is the most applied regression technique which addresses the risk of event time. Thus, the time to first marriage is fitted to the Cox model considering some potential covariates which are thought to explain the age of first marriage to estimate the relative risks. The median duration of singlehood is computed using the non-parametric K-M method.

Kaplan-Meier estimator of survival probability at time t is given by

$$\hat{S}(t) = \prod_{t_i < t} \frac{r_{t_i} - d_{t_i}}{r_{t_i}}$$

where r_{t_i} is the number of risk of experiencing the event at the time t_i, and d_{t_i} is the number of events at that time, with the convention that $\hat{S}(t) = 1$ if $t < t_i$.

Using Greenwood's formula for the variance of survival function

$$\hat{V}(\hat{S}(t)) = (\hat{S}(t))^2 \sum_{r_{t_i}} \frac{r_{t_i}}{r_{t_i} - d_{t_i}}$$

$$\hat{V}[\ln(\hat{S}(t)] = \sum_{r_{t_i}} \frac{r_{t_i}}{r_{t_i} - d_{t_i}}$$

The end point of $100(1 - \alpha)\%$ confidence interval for $S(t)$ on cumulative hazard or log-

Table 1: Variables and categories with sample size

Variables	Categories	Sample size Women	Men
State	Arunachal Predesh	13667	2109
	Assam	27089	4138
	Manipur	12956	1856
	Meghalaya	8662	1220
	Mizoram	11115	1661
	Nagaland	10275	1567
	Sikkim	5114	860
	Tripura	4553	871
Type of place ofresidence	Urban	24754	3923
	Rural	68567	10357
Educational level	No Education	16418	1672
	Primary	13257	2142
	Secondary	54803	8629
	Higher	8843	1837
Religion	Hindu	34828	5512
	Muslim	10443	1500
	Christian	38789	5800
	Others	9266	1468
Ethnicity	SC	6707	1098
	ST	51621	7985
	OBC	12845	1993
	Others	14345	1933
Wealth of family	Poor	40363	6081
	Middle	24945	3891
	Rich	29613	4308
Working status	No	11417	3248
	Yes	3776	11032
Exposed to Media	No	16651	1378
	Yes	76670	12902

Working status and wealth of   family are defined in the note.

survival scale is

$$\exp(\ln \hat{S}(t)) \pm z_{1-\frac{\alpha}{2}} \hat{se}(\ln \hat{S}(t)) \tag{1}$$

We have also computed the estimates of sample median and 95% confidence intervals for all the estimates. The median duration of singlehood is obtained at the time point at which $S(t)$ is less than 0.5. Using the test for Mantel (1966) Log rank test is used to compare the survival experience of duration of singlehood among the categories defined by socio-economic covariates. The regression model for the hazard function that addresses the

study goal is

$$h(t|x) = h_0(t).r(\beta'x) \tag{2}$$

Where $h(t|x)$ is the hazard function, $h_0(t)$ is the baseline hazard and β is the vector of regression parameters and x is the vector of explanatory covariates. Under the model in (2), the ratio of the hazard functions for two individuals (or group of individuals) with covariates x_1 and x_2 is

$$HR(t|x_1, x_2) = \frac{h(t|x_1)}{h(t|x_2)} = \frac{r(\beta'x_1)}{r(\beta'x_2)} \tag{3}$$

From (3), see that if the hazard ratio is easily interpreted then baseline hazard is of little importance. Cox (1972) proposed that the conditional hazard $h(t|x)$ be modelled as the product of $h_0(t)$ and an exponential function which is linear in x that is $r(\beta'x) = e^{\beta'x}$ so that

$$h(t|x) = h_0(t).e^{\beta'x} \tag{4}$$

Under the Cox model in (4), the hazard ratio

$$HR(t|x_1, x_2) = \exp(\beta'x_1 - x_2)$$

As the method in (4) forces the hazard ratio between two individuals to be constant over time, we call it proportional hazards model.

3.   Results

The estimated median duration of singlehood along with 95% confidence interval (C.I.) using (1) for both women and men computed using the non-parametric Kaplan-Meier method is presented in Table 2. The median duration of singlehood for the whole North-East women and men are 21 years and 26 years respectively. Among the eight states Assam and Tripura have the shortest duration of singlehood estimated at 19 years each and Manipur have the longest duration estimated at 23 years for women. For men, Manipur and Nagaland have longest median duration for singlehood of 27 years and Arunachal, Meghalaya and Mizoram have the least median duration of 25 years. Those women and men who are living in urban area have longer median duration of singlehood as compared to their rural counterparts by two years. Women who are educated upto secondary or higher have longer singlehood duration than those women who are educated upto primary or illiterate. Women belonging to Christian and Others religious groups have longer median duration of singlehood. However, those men who are in Hindu religion have longest median duration as compared to the remaining groups. Results show that Muslim women and men have lowest median duration of 18 years and 25 years respectively. Those men who are belonging to SC, ST and OBC category have same median duration of 26 years and others category have the highest (27 years) median duration. Women belonging to SC category have shortest duration of singlehood. Median duration of singlehood for men living in poor and middle wealth categories increases successively by one year. Rich men tend to have longer duration of singlehood (28 years). Women from poor family have shortest singlehood duration among all wealth categories. Generally individuals from richest family have to stay longer in single status. Furthermore, working women have the longer length of singlehood duration than their non-working counterparts but it is contrast in men category. Exposure to mass media is also one of the significant covariates for the study of singlehood duration. Those women

Table 2: Median duration of singlehood for women and men and its *p*-values for testing significance of survivorship experience among categories

Variables	Categories	Median(95% C.I.)		Log rank test (*p*-value)
		Women	Men	
State	Overall	21(20.95,21.05)	26(25.85,26.16)	
	Arunachal Pradesh	20(19.88,20.12)	25(24.61,25.38)	
	Assam	19(18.92,19.08)	26(25.71,26.28)	
	Manipur	23(22.83,23.16)	27(26.56,27.44)	
	Meghalaya	21(20.82,21.17)	25(24.57,25.42)	0.00
	Mizoram	22(21.83,22.17)	25(24.57,25.42)	
	Nagaland	22(21.82,22.18)	27(26.53,27.47)	
	Sikkim	22(21.76,22.24)	26(25.37,26.65)	
	Tripura	19(18.83,19.17)	26(25.84,26.54)	
Type of place of residence	Urban	23(22.87,22.13)	27(26.86,27.31)	0.00
	Rural	20(19.94,20.06)	25(24.82,25.18)	
Educational level	No Education	18(17.91,18.08)	24(23.64,24.35)	
	Primary	19(18.91,19.08)	24(23.67,24.32)	0.00
	Secondary	21(20.94,21.06)	26(25.79,26.20)	
	Higher	28(27.76,28.23)	30(29.59,30.40)	
Religion	Hindu	21(20.92,21.08)	27(26.75,27.25)	
	Muslim	18(17.90,18.10)	25(24.57,25.42)	0.00
	Christian	22(21.94,22.08)	26(25.76,26.23)	
	Others	21(20.82,21.17)	25(24.53,25.46)	
Ethnicity	SC	20(19.82,20.17)	26(25.24,26.57)	
	ST	21(20.92,21.07)	26(25.79,26.20)	0.00
	OBC	21(20.86,21.13)	26(25.59,26.40)	
	Others	21(20.86,21.13)	27(26.54,27.45)	
Wealth of family	Poor	19(18.94,19.06)	25(24.79,25.20)	
	Middle	21(20.90,21.09)	26(25.69,26.30)	0.00
	Rich	23(22.88,23.11)	28(27.70,28.29)	
Working status	No	20(19.86,20.13)	29(28.17,29.80)	0.00
	Yes	22(21.69,22.31)	25(24.84,25.17)	
Exposed to Media	No	19(18.91,19.08)	25(24.62,25.37)	0.00
	Yes	21(20.94,21.05)	26(25.83,26.16)	

who are exposed to mass media have 2 years longer singlehood duration than women who are not exposed to mass media. The last column of Table 2 gives the *p*-values for testing the significant difference of the survival experience among the groups or categories defined by the socio-economic covariates. All covariates are significant at 5% level in the log rank test which in turn suggests that these covariates are important to influence singlehood duration and are potential candidates for the hazards regression model.
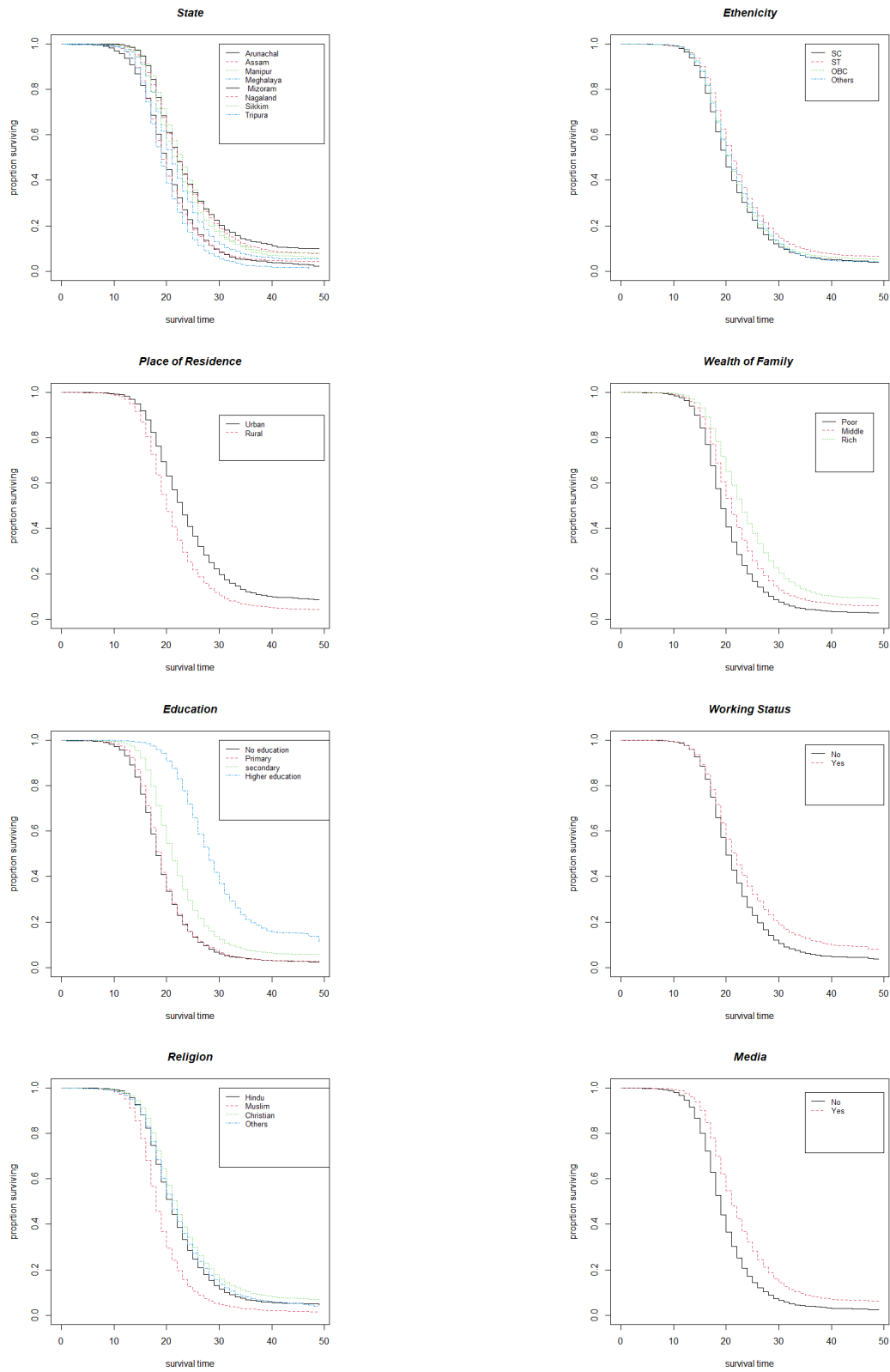
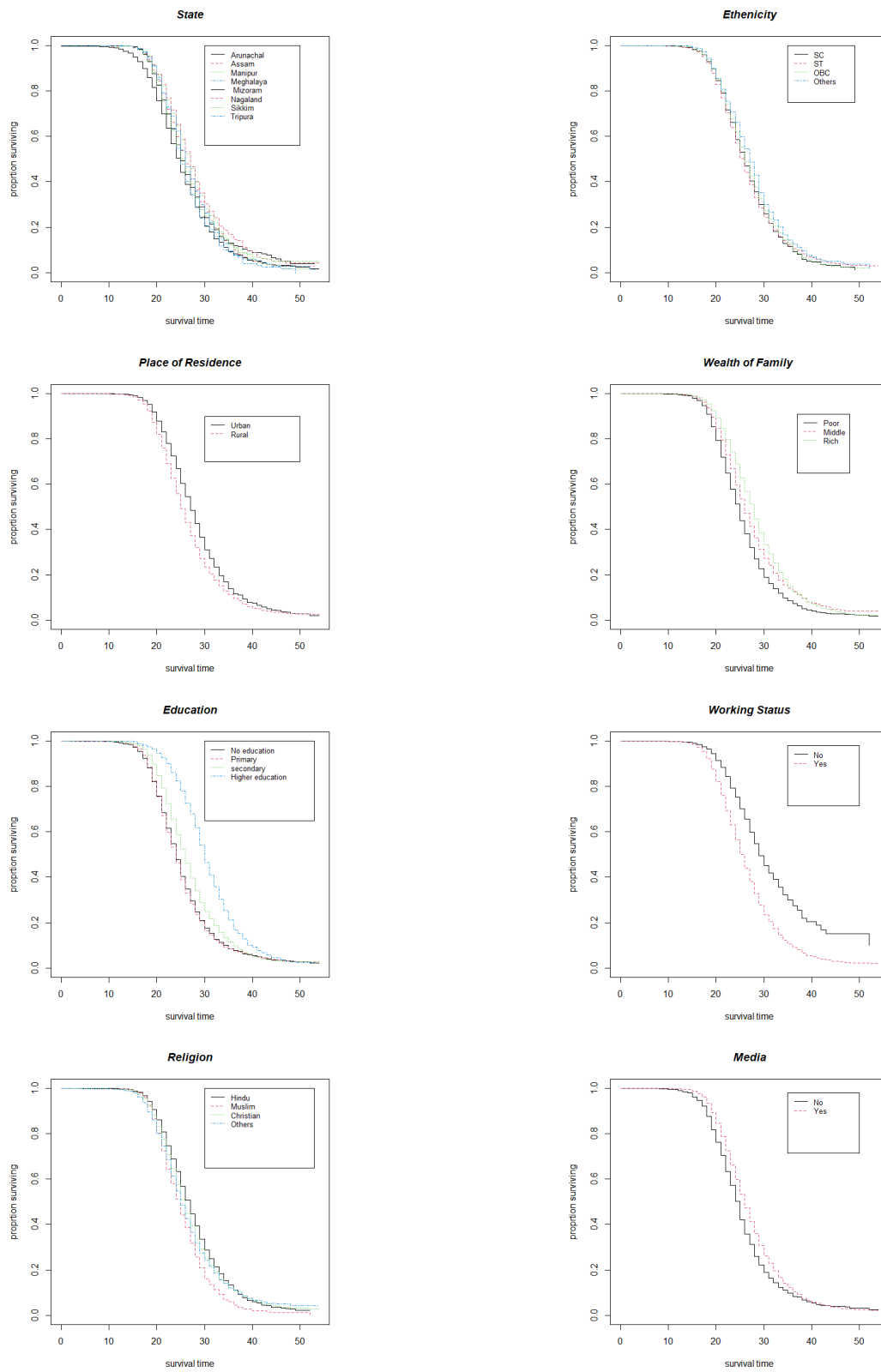Figure 1: Survival curves by background characteristics(women)

Figure 2: Survival curves by background characteristics(men)

3.1.   Results of fitting the Cox hazards regression model

All covariates which are significant (at 5% level) in the bivariate analysis (Table 2) are potential covariates to include in the hazards model as explanatory variables. Consequently, two hazards models one for women and one for men are fitted with the significant covariates to regress the duration of singlehood. The results of fitting the two models are presented in Tables 3 and 4 which include the estimated coefficients, hazards ratio (indicating the reference category), standard error of the estimates and p-values for Wald test for testing the significance of the coefficients. The two models are checked and verified for violation of proportionality assumptions and leverage for influencing observations. Both the models pass the test for proportionality assumption and no influencing observation is present in the data. We discuss it in the next subsection model diagnostics.

3.2.   Model diagnostics

Model based inferences depend completely on the fitted Statistical model. For these inferences to be valid in the real sense of the world, fitted model must provide an adequate summary of the data upon which it is based. A complete and thorough examination of the model's fit and adherence to model assumptions is just as important as careful model development. The methods for assessment of a fitted proportional hazards model essentially consists of i) methods for testing the assumption of proportional hazards and ii) subject specific diagnostic statistics that extend the notion of leverage and influence to the proportional hazards model.

A large number of tests of proportionality assumptions are found in the literature. However, works developed by Grambsch and Therneau (1994) and simulation work by Ng'andu (1997) have shown that an easily performed statistical test and an associated graph yield a powerful and effective method for examining the proportionality assumption. These are the two steps: 1) add the covariates by log time interaction to the model and assess their significance using partial likelihood ratio test, score test or Wald test, and 2) plot the scaled and smoothed scaled Schoenfeld residuals obtained from the model without interaction terms. The result of the two steps should support each other.

The plot for scaled Schoenfeld residuals for some of the covariates are shown in Annexure (Figure 3 and Figure 4) for female and male respectively. Results of statistical tests for proportionality assumptions are shown in Annexure (Table 5 and Table 6) for women and men respectively. We examine the p-values for Wald tests in the interaction terms which are all insignificant suggesting that the covariates have passed the proportionality tests. The graphical plots are more or less flat in all the covariates which support that they all have approximately zero slopes. Covariates with p-value significant at 5% level of significance are removed from the model fit as they may violate proportionality assumption. In the women data all covariates have insignificant p-values and they are all included in the main effects model. All covariates except education are insignificant at 5% level of significance in the men data. So, the final main effects model in the men data include the covariates state, place of residence, religion, ethnicity, wealth of family, working status and media.

It is also important to check outliers by visualising the deviance residuals to identify the influential subjects in the data. The plots of the deviance residuals are shown in Appendix (Figure 5). From the plots it is evident that there are no widely deviant observations in both

the women and men data.

Table 3: Estimated regression coefficients (β), hazards ratio (HR), standard error(SE), two tailed p-values for the proportional hazard model for women

Variables	Categories	β	HR	SE	p-value
State	Arunachal Pradesh(Ref)	-	-	-	-
	Assam	-0.13	0.88	0.04	0.001***
	Manipur	-0.34	0.71	0.04	0.000***
	Meghalaya	-0.17	0.84	0.045	0.000***
	Mizoram	-0.35	0.70	0.045	0.000***
	Nagaland	-0.35	0.70	0.04	0.000***
	Sikkim	-0.21	0.81	0.05	0.000***
	Tripura	0.18	1.20	0.05	0.000***
Type of place of residence	Urban(Ref)	-	-	-	-
	Rural	0.07	1.08	0.03	0.005***
Educational level	No Education(Ref)	-	-	-	-
	Primary	0.01	1.01	0.03	0.725
	Secondary	-0.38	0.69	0.03	0.000***
	Higher	-1.25	0.29	0.05	0.000***
Religion	Hindu(Ref)	-	-	-	-
	Muslim	0.35	1.41	0.05	0.000***
	Christian	0.06	1.06	0.04	0.111
	Others	-0.03	0.97	0.04	0.458
Ethnicity	SC(Ref)	-	-	-	-
	ST	-0.14	0.87	0.04	0.001***
	OBC	-0.06	0.94	0.04	0.151
	Others	-0.02	0.97	0.04	0.529
Wealth of family	Poor(Ref)	-	-	-	-
	Middle	-0.69	0.94	0.03	0.019**
	Rich	-0.09	0.92	0.03	0.006***
Working status	No(Ref)	-	-	-	-
	Yes	-0.08	0.93	0.02	0.001***
Media	No(Ref)	-	-	-	-
	Yes	-0.004	1.00	0.03	0.888
Ref=reference,	***=p <0.01, **=p <0.05.				

3.3. Interpretation of fitted models

The popularity of a fitted regression hazards model is due to its ease in interpreting and understanding the hazards ratios which literally gives relative risk of experiencing the event of interest with respect to a reference category for a categorical covariate. The 4th column of Table 3 present the relative risk of first marriage for women and men respectively,

Table 4: Estimated regression coefficients (β), hazards ratio (HR), standard error(SE), two tailed p-values for the proportional hazard model for men

Variables	Categories	β	HR	SE	p-value
	Arunachal Pradesh(Ref)	-	-	-	-
	Assam	-0.37	0.69	0.04	0.001***
	Manipur	-0.36	0.70	0.05	0.000***
State	Meghalaya	-0.20	0.82	0.05	0.000***
	Mizoram	-0.18	0.83	0.05	0.000***
	Nagaland	-0.44	0.64	0.05	0.000***
	Sikkim	-0.05	0.95	0.05	0.334
	Tripura	-0.25	0.78	0.06	0.000***
Type of place	Urban(Ref)	-	-	-	-
of residence	Rural	0.05	1.05	0.03	0.078*
	No Education(Ref)	-	-	-	-
Educational level	Primary	0.01	1.01	0.03	0.725
	Secondary	-0.38	0.69	0.03	0.000***
	Higher	-1.25	0.29	0.05	0.000***
	Hindu(Ref)	-	-	-	-
Religion	Muslim	0.20	1.22	0.06	0.000***
	Christian	0.06	1.06	0.04	0.171
	Others	0.013	1.01	0.05	0.795
	SC(Ref)	-	-	-	-
Ethnicity	ST	-0.03	0.97	0.05	0.540
	OBC	-0.007	0.99	0.05	0.822
	Others	-0102	0.88	0.05	0.014**
	Poor(Ref)	-	-	-	-
Wealth of family	Middle	-0.24	0.79	0.03	0.000***
	Rich	-0.37	0.69	0.03	0.000***
Working status	No(Ref)	-	-	-	-
	Yes	0.65	1.91	0.04	0.000***
Media	No(Ref)	-	-	-	-
	Yes	0.005	1.01	0.04	0.884
Ref=reference,	***=$p <0.01$, **=$p <0.05$, *=$p <0.1$				

along with the regression coefficients and standard error of coefficients for different socio-economic covariates. In Table 3 and Table 4, the hazards ratios (HR) are shown for all the North East states (women) with Arunachal as the reference category state. From the p-values (Wald test)in the last column of Table 3, it is evident that the HR for all states are significant at 5% level. Manipur's HR of 0.71 reveals that women in Manipur have nearly 29% less risk of first marriage as compared to women in Arunachal Pradesh. Similarly, women in Nagaland have significantly lower risk (30%) of first marriage as compared to Arunachal women. However, women in Tripura marry earlier as the HR of 1.2 indicates that the risk of first marriage for women in Tripura is nearly 1.2 times that of Arunachal women.

As of now the insignificant HR's may not be interpreted as such. For men (Table 4) all the HRs for all the states except Sikkim are highly significant. Literally, the risks of marriage for men in these states are significantly lower than that of Arunachal men.

In the whole North East region, women who live in urban area have lower risk of age at marriage. Rural women have 1.08 times higher risk of first marriage as compared to urban women. Approximately, men in rural area have the same higher risk (5%) of first marriage as compared to urban men. Among women who are educated upto primary or no education the risk of marriage does not differ significantly. However, those women who have education upto secondary and higher have significantly lower risk of first marriage upto the tune of 31% and 70% respectively as compared women who have no education at all. Among the religious groups at the community level, the relative risk of first marriage for Muslim women is significantly higher than the Hindu women. Muslim women have 51% higher chance to marry earlier than the Hindu women. Similarly, for Muslim men have 22% more chance of first marriage as compared to men in the Hindu religion. Other categories of religion are not significant. Among the ethnic groups, ST category has HR = 0.97 which interpret that women of ST category have 3% less likely to marry as compared to women belonging to SC category. Other categories of ethnicity are not significant.

At the household level, women living in middle and rich family exhibit lower risks of marriage as compared to women in poor family. Women in middle and rich wealth quintiles are 6% and 8% less likely to marry as compared to women in poor wealth quintile respectively. Similarly, men belonging to middle and rich wealth quintiles are respectively 21% and 31% less likely to marry as compared to men belonging to poor wealth quintile.

At the individual level, working status of both women and men has significant effect on singlehood duration. Working women have less chance of marriage to the tune of 7% less as compared to women living with no working status. However, for men the result is just the reverse as working men have more chance to marry to the tune of 1.9 times more likely as compared to non-working men.

4.   Discussion

Age at marriage is one of the significant life events for every individual. It signals the entry of each individual into the state of being married. This study attempted to investigate the median duration of singlehood for North east India using the NFHS-4 data. Cox proportional hazards model is fitted to assess the significant effect of various covariates on the singlehood duration.

First, it is observed that the duration of singlehood varies among groups of population identified by different covariates. In NFHS-4 (2017), 2015-16 (International Institute for Population Sciences (IIPS) and ICF, 2017) the median age at marriage for women and men in India is estimated to be 18.6 years and 24.5 years respectively. However, for North–east region the median age at first marriage for women is 21 years and 26 years for men. This indicates that the people in North east India are more likely to live in single status than the people in the rest of the country. North East region of India comprises of eight states with a different socio-cultural set up from the mainland India. From the results states with Christian as main religion like Arunachal, Meghalaya and Mizoram have least duration of singlehood in men category. Whereas Manipur and Nagaland show higher estimates of

median duration of singlehood than others in men category.

Rural women and men are more likely to get married at early ages which in turn indicate that urban people have longer median duration of singlehood as compared to rural people in the North east region. Educational level of individuals is one of the important determinant factors for the early marriage as many literatures have cited. North East women and men with higher education have longer duration of singlehood than others with low educational level which is in line with the findings of other studies. The chance of singlehood for women increases with increase in educational level.

For the whole country according to NFHS-4 (2017), 2015-16 (International Institute for Population Sciences (IIPS) and ICF, 2017) and NFHS-3 (2007), 2005-06 (International Institute for Population Sciences (IIPS) and Macro International, 2007) reports, Hindus and Muslims have similar median age at marriage for the whole country. However, in the North East region, Muslims are more likely to marry at early ages than the Hindus and Christians. Thus, Muslims have shorter duration of singlehood.

Present paper also explores the effect of ethnicity on the survival experience on the duration of singlehood. Schedule tribe women population has lower risk of marriage as compared to Schedule caste women population in the region less chance of marriage than the schedule caste women. Besides, men from others ethnicity groups are less likely to get married than those in other categories. Another important finding is that wealth of the family significantly affects the duration of singlehood. Women from poor family are more likely to marry earlier than others. In a similar manner,men fro m poor family have higher chance of marriage than men from richer families. Last but not the least; we assess the influence of working status on the singlehood duration. Working women are more likely to be in single state than non-working women. Interestingly, this phenomenon is just the reverse for men, which shows that men who are currently working have more risk of marrying earlier than their non- working counterparts. This is also in line with some of the findings in the literature. Exposure to mass media has no significant effect on the study of duration of singlehood in North East region.

5.   Conclusion

Marriage is a major life event which basically marks the onset of married couples contributing to human reproduction. As such marriage is considered as an important proximate determinant of fertility for a country or a region. As early marriages are expected to contribute more births it is important to increase the age at marriage of both men and women in order to reduce fertility NFHS-3 (2007). The median age at marriage in India increases from NFHS-3 (2007) to NFHS-4 (2017) by two years for both men and women. In order to further improve the age at marriage, the policy makers have to give further attention to the socio-economic disparities of age at marriage in the country. Regional findings will be helpful in the present context and as such the findings in this paper could be helpful to policy and programme planners while addressing the issue of population control through improvement in age at marriage.

Note

1. According to NFHS-4 (2017), Wealth index is a measure of living standards based on households' ownership of items such as televisions to housing features such as drinking water sources. The population is divided into five equally sized groups based on index. The top 20% form the richest, and the bottom 20% the poorest quintile. In the present analysis the wealth index is condensed into 3 categories *viz* (1) poorer and poorest into poor, (2) middle and (3) richer and richest to rich.

2. Women who are currently working outside the home for earning are considered as working women. Such working women are categorized as "Yes" otherwise "No".

Acknowledgements

References

Bongaarts, J. (1978). A framework for analyzing the proximate determinants of fertility. *Population and Development Review*, **34**, 105–132.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **34**, 187–202.

Gangadharan, L. and Maitra, P. (2000). *The Effect of Education on the Timing of Marriage and First Conception in Pakistan*. Monash University.

Grambsch, P. M. and Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, **81**, 515–526.

Hayase, Y. and Liaw, K.-L. (1997). Factors on polygamy in Sub-Saharan Africa: findings based on the demographic and health surveys. *The Developing Economies*, **35**, 293–327.

Jensen, R. and Thornton, R. (2003). Early female marriage in the developing world. *Gender & Development*, **11**, 9–19.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.

Kumar, S. (2016). Measuring child marriage from census and large scale data systems in india. *Demogr India*, **45**, 59–76.

Kumchulesi, G., Palamuleni, M., and Kalule-Sabiti, I. (2011). Factors affecting age at first marriage in malawi. In *Sixth African Population Conference, Ouagadougou-Burkina Faso*, pages 5–9.

Lalmalsawmzauva, K., Nayak, D., and Vanlalvena, R. (2011). Interface between developments and female age at marriage in mizoram. *Transaction, Journal of the Institute of Indian Geographers*, **33**, 139–149.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep*, **50**, 163–170.

Matlabi, H., Rasouli, A., Behtash, H. H., Dastjerd, A. F., and Khazemi, B. (2013). Factors responsible for early and forced marriage in iran. *Science Journal of Public Health*, **1**, 227–229.

McLaughlin, D. K., Lichter, D. T., and Johnston, G. M. (1993). Some women marry young: Transitions to first marriage in metropolitan and nonmetropolitan areas. *Journal of Marriage and the Family*, **55**, 827–838.

NFHS-3 (2007). *National Family Health Survey (NFHS-3), 2005-06: India (2 v.+ suppl.)*, volume 1. International Institute for Population Sciences.

NFHS-4 (2017). *National Family Health Survey (NFHS-4), 2015-16: India*, volume 1. International Institute for Population Sciences.

Ng'andu, N. H. (1997). An empirical comparison of statistical tests for assessing the proportional hazards assumption of cox's model. *Statistics in Medicine*, **16**, 611–626.

Spreitzer, E. and Riley, L. E. (1974). Factors associated with singlehood. *Journal of Marriage and the Family*, **36**, 533–542.

Stein, P. J. (1975). Singlehood: An alternative to marriage. *Family Coordinator*, **24**, 489–503.

Weinberger, M. B. (1987). The relationship between women's education and fertility: selected findings from the world fertility surveys. *International Family Planning Perspectives*, **13**, 35–46.

Westoff, C. F. and ORC Macro, C. (2003). Trends in marriage and early childbearing in developing countries. **5**.

ANNEXURE

Table 5: Test for proportional hazards assumption(women)

| Covariants	Categories	Chisq	*p*-value
State	Assam	0.120	0.729
	Manipur	0.517	0.472
	Meghalaya	0.630	0.428
	Mizoram	0.222	0.638
	Nagaland	0.268	0.604
	Sikkim	0.533	0.465
	Tripura	1.170	0.279
Type of place of residence	Rural	1.950	0.163
Educational level	Primary	2.600	0.107
	Secondary	0.001	0.969
	Higher	119	0.000
Religion	Muslim	1.880	0.171
	Christian	0.518	0.472
	Others	0.206	0.650
Ethnicity	ST	0.168	0.682
	OBC	3.060	0.081
	Others	1.290	0.257
Wealth of family	Middle	4.430	0.035
	Rich	15.600	0.000
Working status	Yes	16.600	0.000
Media	Yes	3.710	0.054
State:time	Assam	1.300	0.254
	Manipur	2.430	0.119
	Meghalaya	0.037	0.847
	Mizoram	1.420	0.234
	Nagaland	0.739	0.390
	Sikkim	0.365	0.546
	Tripura	0.075	0.784
Type of place of residence :time	Rural	0.592	0.442
Educational level :time	Primary	2.600	0.107
	Secondary	0.001	0.969
	Higher	119	0.000
Religion:time	Muslim	1.880	0.171
	Christian	0.518	0.472
	Others	0.206	0.650
Ethnicity :time	ST	0.168	0.682
	OBC	3.060	0.081
	Others	1.290	0.257
Wealth of family :time	Middle	0.757	0.384
	Rich	0.180	0.672
Working status:time	Yes	3.380	0.066
Media :time	Yes	0.483	0.487
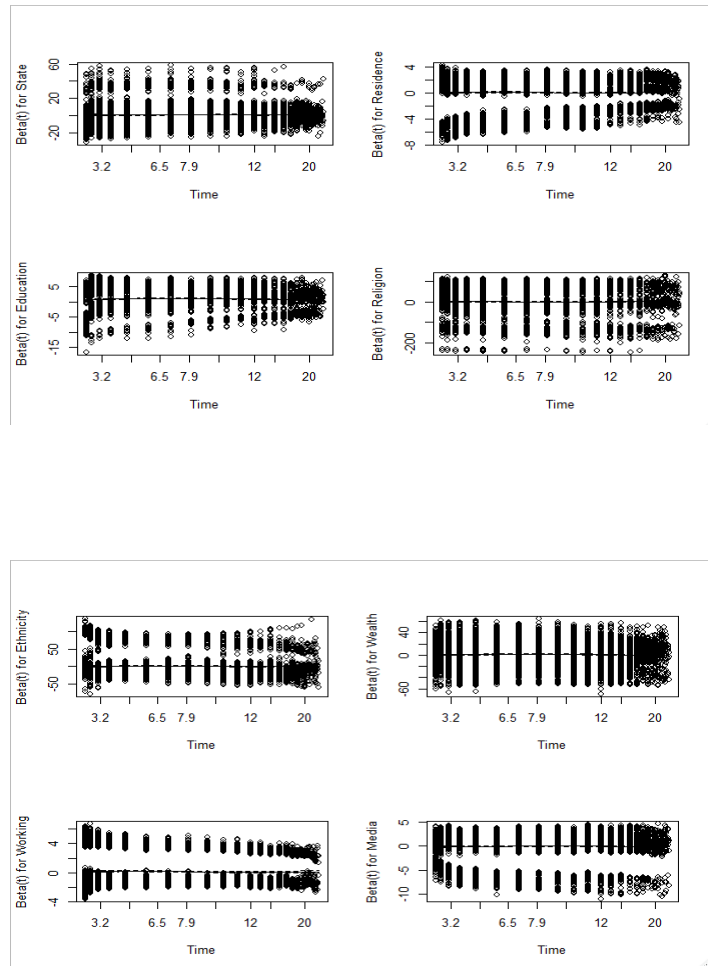
Figure 3: Plot of scaled Schoenfeld residuals and smoothed scaled Schoenfeld residuals for assessing proportionality assumptions in some covariates (women)

Table 6: Test for proportional hazards assumption(men)

| Covariates	Categories	Chisq	p-value
State	Assam	0.800	0.371
	Manipur	1.450	0.229
	Meghalaya	8.640	0.003
	Mizoram	4.330	0.037
	Nagaland	0.137	0.711
	Sikkim	0.172	0.678
	Tripura	0.143	0.706
Type of place of residence	Rural	4.840	0.028
Educational level	Primary	6.400	0.011
	Secondary	1.510	0.219
	Higher	119.000	0.000
Religion	Muslim	0.096	0.757
	Christian	1.310	0.253
	Others	0.051	0.822
Ethnicity	ST	0.713	0.398
	OBC	0.263	0.608
	Others	0.000	0.922
Wealth of family	Middle	0.080	0.777
	Rich	0.732	0.392
Working status	Yes	1.400	0.237
Media	Yes	0.035	0.852
State:time	Assam	1.720	0.190
	Manipur	0.657	0.418
	Meghalaya	8.920	0.003
	Mizoram	6.340	0.012
	Nagaland	0.807	0.369
	Sikkim	0.102	0.749
	Tripura	0.196	0.658
Type of place of residence :time	Rural	1.890	0.169
Educational level :time	Primary	7.010	0.008
	Secondary	7.480	0.006
	Higher	91.700	0.000
Religion:time	Muslim	0.151	0.698
	Christian	0.645	0.422
	Others	0.123	0.725
Ethnicity :time	ST	0.183	0.669
	OBC	0.137	0.711
	Others	0.082	0.775
Wealth of family :time	Middle	0.070	0.792
	Rich	0.711	0.399
Working status:time	Yes	0.676	0.411
Media :time	Yes	0.013	0.911
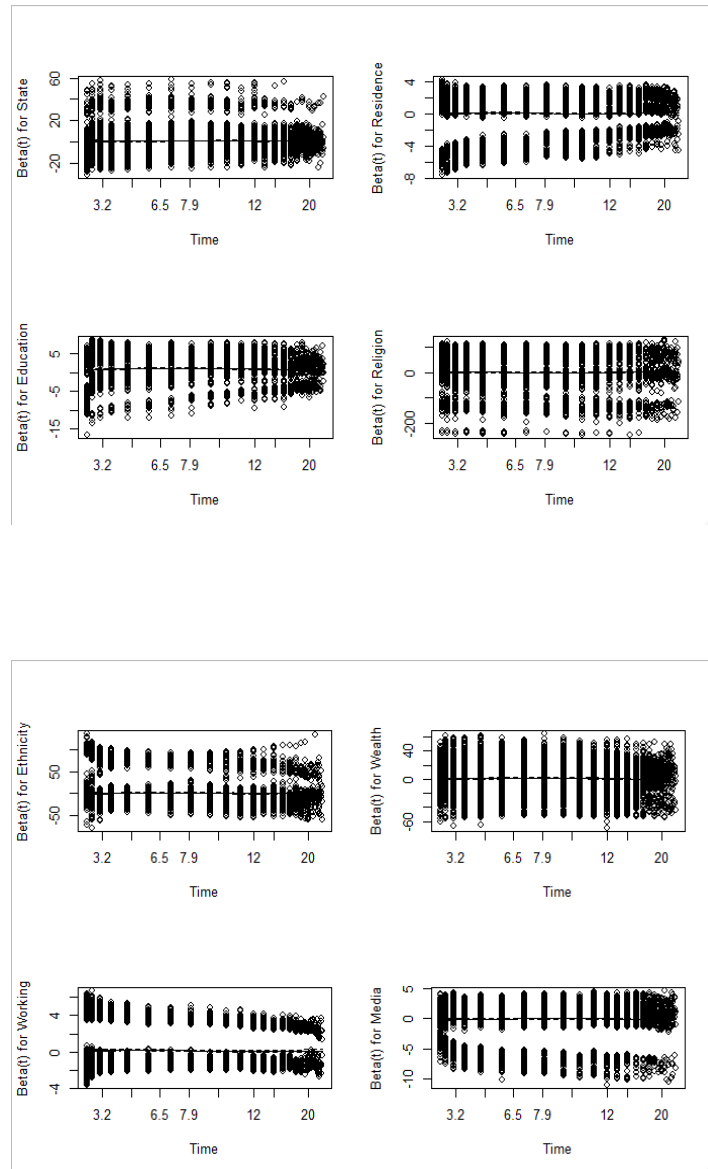
Figure 4: Plot of scaled Schoenfeld residuals and smoothed scaled Schoenfeld residuals for assessing proportionality assumptions in some covariates (men)
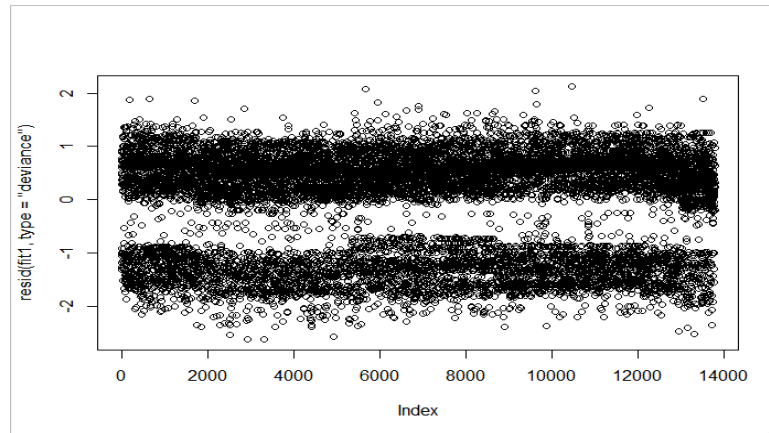
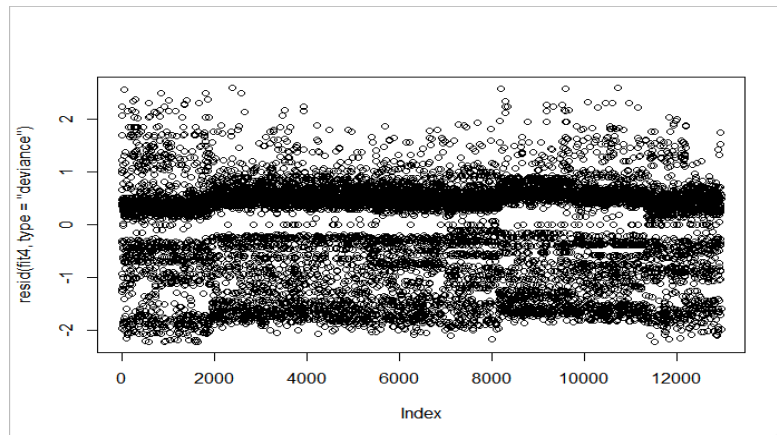**Figure 5a: Deviance residuals for women covariates**

Figure 5b: Deviance residuals for men covariates

Figure 5: Deviance residuals