

Socioeconomic and Demographic Determinants of Birth Interval in Bangladesh: A Statistical Analysis Using Survival Mixed Regression Model and Proportional Hazard Regression Model

Sharlene Alauddin¹, Md. Akhtarul Islam² and M. Ershadul Haque³

^{1,2}Statistics Discipline, Khulna University, Khulna 9208, Bangladesh
email: sharlene@ku.ac.bd, akhtarulstat@gmail.com

³Department of Statistics, Dhaka University, Dhaka 1000, Bangladesh
email: ershad_sbi@du.ac.bd

Received 28 February 2019; accepted 29 March 2019

Abstract. To examine the unobserved covariate's influence on the distribution of pace of childbearing is considered as a great important factor for understanding the insights of birth interval patterns, maternal and childhealth during a woman's whole reproductive period. The paper examines the impact of birth interval considering relative influencing factors. The application of Survival Mixed Regression Model considers the existence of such unobserved covariates that influences the length of birth interval. The objective of this study is to model the length of time to preceding birth interval within the reproductive health of mothers and to identify the socioeconomic and demographic factors that cause variation in the length of birth interval. The data extracted from BDHS 2014, conducted under two-stage stratified sampling design were utilized for the purpose of analysis. A Survival Mixed Regression model was utilized in this study and several factors were investigated for birth interval by fitting this model. A semi-parametric model (Cox Proportional Hazard Regression model) was also employed for comparison purpose. In Survival Mixed Regression Model single mother with multiple birth interval are considered to be a cluster. The estimated parameters of covariates is assessed and compared under these two statistical models. The result of this study indicates that the unobserved cluster effect has a sizable impact on birth interval in Bangladesh.

Keywords: Cox proportional hazard model, survival mixed regression model, frailty

AMS Mathematics Subject Classification (2010): 91Cxx

1. Introduction

Birth interval measured by the number of months between the birth of the child under study and the immediately succeeding birth to the mother is a susceptible analysis method for measuring mechanism of fertility [1, 2]. Number of births during a woman's whole reproductive period depends on the duration of intervals which are associated with some factors. Here in this study, we examine the determinants of birth interval by considering two models such as Cox PH model and Survival Mixed Regression (SMR)

Model. Impact, consequences and differences between these two models are represented here. The conventional Cox's PH model assumes that the subjects of interest under given experimental conditions are independent and identically distributed and hence homogeneous by nature. But sometimes it is necessary to modify the homogeneity assumption in the conventional Cox's model in order to account the heterogeneity fact [3]. Thus to remove the biasness of Cox PH model a "hidden heterogeneity" or "frailty" or "random effect" term is included. Then this modified model is known as Survival Mixed Regression model. In survival analysis, the importance of survival mixed regression model is enormous. Heckman have argued that results obtained from time to event models can be misleading unless unobserved heterogeneity has been considered [4-6]. Similarly, Trussell and Rodriguez (1990) noted that the failure to correct for unobserved heterogeneity can lead to hazards that either decline steeply or rise slowly than the true hazard which is resulting in biased parameter estimates [7]. Survival Mixed Regression model assumes that the unmeasured covariates simultaneously adjust within any cluster that consider heterogeneity [8]. Thus this Frailty term follows a distribution which plays a double role in describing both the non-proportionality and the intra class correlation even after considering censoring status [9]. In this study, Survival Mixed Regression model is used simply to explore the correlation among the observations of groups. Here children of each mother are considered to share the same frailty. It is expected that birth interval among children within each mother to be correlated because birth interval pattern of one mother would inherently be more different than others.

2. Data and methods

This study is accomplished and evaluated by using the nationwide well representative survey called Bangladesh Demographic and Health Survey (BDHS 2014) which is a very valuable source of national and divisional information [10]. BDHS data is a monitoring indicators in population, health, and nutrition studies which is mainly funded by US Agency for International Development (USAID) and is accomplished under the authority of the National Institute of Population Research and Training (NIPORT) of the Ministry of Health and Family welfare, Bangladesh. The BDHS survey was implemented through a private research firm located in Dhaka named Mitra and Associate. At the first stage of sampling 600 clusters were selected where in the second stage 30 household were selected in a systematic way. This is formally known as second stage of sampling where a total of 17,989 residential households are selected where 18,245 ever-married women age 15-49 were interviewed. The numbers of eligible women were interviewed where 6,324 are from urban areas and 11,696 are from rural areas. To collect a detailed history on birth interval, the required information on socioeconomic and demographic factors were obtained from the children of ten years with the resulting data set of size 6526.

2.1. Dependent variable

In this study birth interval in month is the dependent variable. It is defined as a continuous variable and denoted by "Time". The time variable measured in month which is defined as,

$$\text{Time} = \begin{cases} \text{Succeeding birth interval for the index child (event)} \\ \text{Age of the child at the end of the study (non- event)} \end{cases}$$

Socioeconomic and Demographic Determinants of Birth Interval in Bangladesh

2.2. Explanatory variable

This study deals with a large number of explanatory variables containing Division, Place of Residence, Mother's Education, Wealth Index, Mother's Access to Media, Mother's Working Status, Sex of Child, Survival Status of Child, Mother's Age at Birth.

2.3. Cox proportional hazard model

Let $X = (X_1, X_2, \dots, X_p)$ be the $P \times 1$ vector of covariates which is associated with $P \times 1$ vector of regression coefficients $\beta = (\beta_1, \dots, \beta_p)$. Suppose that T be a survival time in the presence of the covariate and $h_0(t)$ is the hazard function at the time point t in the absence of the covariate which is known as baseline hazard function. Now under the Cox proportional hazard model the hazard function in the presence of covariate can be expressed as,

$$h(t) = h_0(t) c(x)$$

where $c(x)$ is a positive valued function which has a known parametric form such as

$$c(x) = \exp(\beta'x).$$

The principles of the Cox proportional hazard model usually create a link between survival time of an individual and corresponding associated covariates [11]. It explores the relationship between the event of interest and several explanatory variables. Finally, Cox model involves examining the coefficients for each explanatory variable and get the required result. In this study, t_{ij} represents our event of interest; the birth interval corresponding to the j th child of the i th mother of the cluster. Let $X_{ij} = (X_{ij1}, \dots, X_{ijp})'$ is the associated p -dimensional covariate vector ($i = 1, \dots, n; j = 1, \dots, m_i$). The hazard function for the birth interval t_{ij} under the Cox proportional hazard model can be written as,

$$h(t_{ij}) = h_0(t_{ij}) \exp(\beta'x_{ij})$$

2.4. Survival mixed regression model

The Survival Mixed Regression model is a special case of Cox PH model that considers a random effect term. It is a statistical modeling concept along with frailty approach that aims to account for heterogeneity caused by unmeasured factors [12]. Survival Mixed Regression models assume that similar observations share same frailty, even though frailty may vary from group to group. As a result, a major problem in repeatable events thought to be arise from unobserved heterogeneity. To adjust this unobservable correlation a non-negative random term or frailty is used which is represented by w_i . It is mainly used to analyze time to event data and build up a correlation among unobservable cluster specific covariates [13]. Under the Survival Mixed Regression model, the hazard function condition given with a random cluster effect can be expressed as,

$$h(t_{ij}|w_i) = h_0(t_{ij})w_i \exp(\beta'x_{ij})$$

where, in this study w_i is the random effect term corresponding to the i th mother known as frailty. In this paper, we introduce a random effect at the mother (cluster) level and consider each mother as a cluster and its subsequent birth interval as cluster's observation. Significance of the effect of heterogeneity due to mother on birth interval of Bangladeshi women have been discussed in this study.

3. Results

3.1. Univariate analysis

This data contain total 6526 observations and 10 explanatory variables. The results for the variable region represents that Chittagong has the highest frequency of respondent with 18.6%. On the other hand Barisal has the lowest frequency of respondent with 11.0% percentage. For the place of residence variable 65.5% of the respondents are from rural area and remaining 34.5% are from urban areas. In education section, it is observed that 9.4% women have no education, 26.8% women have primary education, 50.8% women have secondary education and only 12.9% women completed their higher education. It shows a large number of respondent (91.0%) are Muslim and only 9.0% of respondent are from other religion. It is also observed that, from 6526 respondent 33.0% have the poor economic status, 33.2% are from middle economic status and 33.8% of respondent have rich economic status. In case of respondent access to media, it is found that 65.7% are exposed and 34.3% are non-exposed to media. It is quiet surprising that a large number of respondent 73.4% are housewives while only 26.6% are working women. This work reveals that the group of mothers age less than 20 possesses the highest frequency with 79.0%. Thus only 21.0% mother age is greater than 20 at birth. Here in the selected data, mothers have 49.4 % female children and 50.6% are male children. A large number of these children (94.1%) are alive and only 5.9% of them are died.

3.2. Bivariate analysis

To observe the association between birth interval and selected socioeconomic and demographic variables, a bivariate analysis is conducted. The survival curve for different variable are plotted using Kaplan-Meier estimator. Also log rank test is used to compare the survival probability between two or more group of individuals. The p-value obtained from this test is also included to assess the significant effect of these variables on birth interval. Survival curve for all the variables are given below but in final model we only considered significant variables.

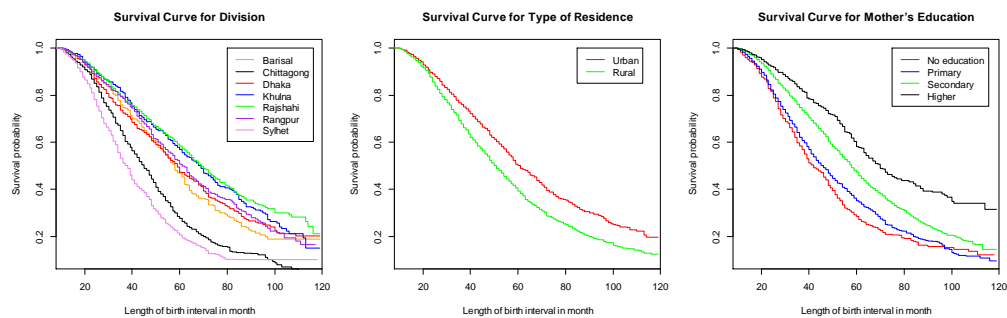


Figure 3.1: log- rank test: p-value = 0.0 < 0.05

Figure 3.2: log- rank test: p-value = 2.11e-15 < 0.05)

Figure 3.3: log- rank test: p-value = 0.0 < 0.0

Socioeconomic and Demographic Determinants of Birth Interval in Bangladesh

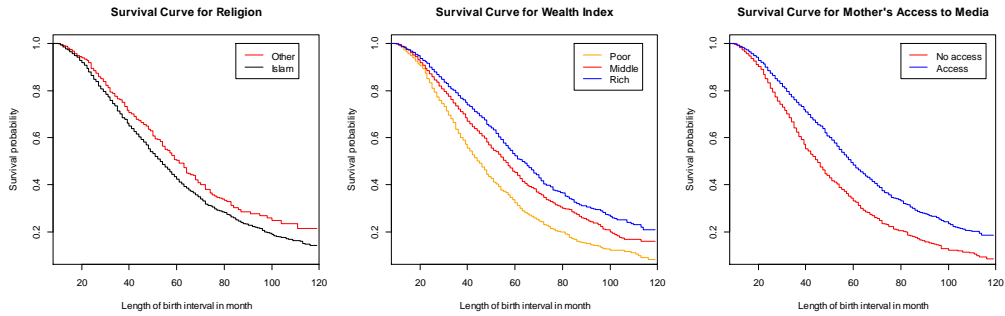


Figure 3.4: log- rank test: p-value = 0.00156 < 0.05

Figure 3.5: log- rank test: p-value = 0.0 < 0.01

Figure 3.6: log- rank test: p-value = 0.0 < 0.01

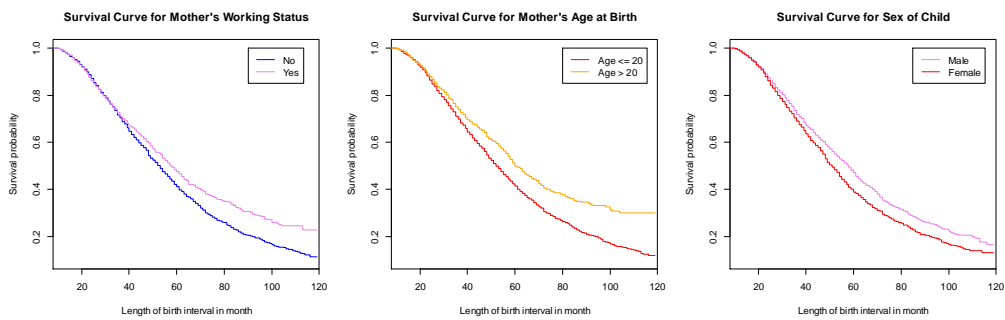


Figure 3.7: log- rank test: p-value = 1.7e-6 < 0.1

Figure 3.8: log- rank test: p-value = 9.2e-10 < .01

Figure 3.9: log- rank test: p-value = 1.9e-06 < .01



Figure 3.10: log- rank test: p-value = 0.0 < 0.01

3.3. Multivariate analysis

Table 1: Estimates of the hazard ratios and corresponding coefficients of the Cox Proportional Hazards model (Model I) and Survival Mixed Regression Model (Model II).

Variables	Category	Model I		Model II	
		β	HR	β	HR
Division	Barisal (RC)	-	-	-	-
	Chittagong	0.495	1.64 (0.0)	0.656	1.927 (0.0)
	Dhaka	0.090	1.09 (0.214)	0.157	1.170 (0.13)
	Khulna	-0.184	0.832 (0.019)	-0.227	0.797 (0.04)
	Rajshahi	-0.231	0.793 (0.003)	-0.313	0.731 (0.005)
	Rangpur	-0.119	0.887 (0.120)	-0.104	0.901 (0.34)
	Sylhet	0.634	1.89 (0.0)	0.889	2.434 (0.0)
Type of residence	Urban (RC)	-	-	-	-
	Rural	0.117	1.125 (0.007)	0.134	1.143 (0.034)
Mother's education	No education (RC)	-	-	-	-
	Primary	0.067	1.069 (0.268)	-0.038	0.963 (0.69)
	Secondary	-0.068	0.934 (0.264)	-0.252	0.778 (0.007)
	Higher	-0.190	0.827 (0.033)	-0.376	0.687 (0.004)
Religion	Others (RC)	-	-	-	-
	Islam	0.175	1.191 (0.008)	0.244	1.276 (0.009)
Wealth index	Poor (RC)	-	-	-	-
	Middle	-0.237	0.789 (0.0)	-0.265	0.767 (0.0)
	Rich	-0.361	0.697 (0.0)	-0.451	0.637 (0.0)
Mother's access to media	No access (RC)	-	-	-	-
	Access	-0.129	0.878 (0.004)	-0.155	0.857 (0.021)
Mother's working status	No (RC)	-	-	-	-
	Yes	-0.131	0.877 (0.001)	-0.076	0.927 (0.20)
Mother's age at birth	Below age 20 (RC)	-	-	-	-
	Above age 20	-0.159	0.853 (0.002)	-0.042	0.959 (0.58)
Sex of child	Male (RC)	-	-	-	-
	Female	0.170	1.186 (0.0)	0.215	1.240 (0.0)
Survival status of previous child	Dead (RC)	-	-	-	-
	Alive	-1.207	0.299 (0.0)	-1.892	0.151 (0.0)

From model II of table 1, the following results were observed. The mother of Chittagong division have $(1.927-1)*100\% = 92.7\%$ higher risk of occurring the next birth compared to the mother of Barisal division. The mothers who are residing in Khulna and Rajshahi

Socioeconomic and Demographic Determinants of Birth Interval in Bangladesh

division have 20.3 % (p-value< 0.05) and 26.8% (p-value < 0.01) lower risk of occurring the next birth. In other words the mothers from Khulna and Rajshahi division take their next child at a higher interval than that of Barisal division. In case of type of residence, the hazard ratio for the mothers who are living in rural area is 1.143. That means, mother residing in rural area (p-value < 0.05) have 14.3% higher risk of occurring the next birth than the mother in urban area. It is expected that highly educated mothers have large birth spacing. The hazard ratio for the mothers who have completed higher education is 0.687. It indicates that mothers with higher education have 31.3% lower risk of occurring the next birth than those of non-educated mothers (p-value < 0.01). Similarly, mothers who passed secondary education have 22.2% (p-value < 0.01) lower risk of occurring the next birth. In case of religion, the risk of occurring next birth is 27.6% higher for Muslim mothers as compared to the other religion indicating that, Muslim mothers take their next child at a short interval. In the case of wealth index, the risk of occurring the next birth have been found 23.3% which is lower for the middle family compared to those of poor families (p-value < 0.01). Similarly, the rich family experiences 36.3% (p-value < 0.01) lower risk of occurring the next birth compared to poor class families. A significant difference on birth interval is observed by mother's access to media. The Table 1 shows that the hazard ratio for this variable is 0.857. Thus the risk of next birth by time for the exposure which shows that mothers who have access to media have 14.4% (p-value<0.05). The variable mother's working status doesn't show a significant difference on birth interval. The hazard ratio for working mothers is 0.927 which is insignificant (p-value>0.01). This finding is contradictory to Cox PH model where the variable found to be significant (p-value<0.01). In case of mother's age at birth, hazard ratio for the mother's age above 20 is 0.959 (p-value> 0.05) which is statistically insignificant, contrast with Cox PH model where this variable is found to be significant. The Table 1 shows a significant difference on birth interval by sex of previous child. The hazard ratio for the mother already having female child is 1.24(p value < 0.01). That means, if the children is female, mothers have 24% higher risk of occurring the next birth compare to male children. A significant difference on birth interval is observed by survival status of the previous child. The Table1 shows that if the children are alive mothers with hazard ratio 0.151 have 84.9% lower risk of occurring the next birth compared to the children who died.Finally, we may conclude that, the variable division, type of residence, mother's education, religion, wealth index, mother's access to media, sex of child and survival status of previous child are found to be significant in both Cox PH model and Survival Mixed Regression model. On the other hand, mother's working status and mother's age at birth gave differnt results. Effects of these variable were found significant on birth interval under Cox PH model. However they have insinificant effects on birth interval in case of Survivsl Mixed Regression model.The results obtained from the Survival Mixed Regression model with the gamma frailty distribution indicate that the unobserved cluster effects have sizeable (individual) impact on birth interval where the variance of random effect is 1.0357 but Cox PH model ignores this term.

Table 2: AIC comparison for Cox PH Model and survival mixed regression model

Model	AIC value	Cluster Variance
Cox PH model	48567.49	-
Survival Mixed Regression model	44290.47	1.035765

AIC difference	4277.02	
----------------	---------	--

From the Table 2 it is observed that Survival Mixed Regression model have smaller AIC value than Cox PH model indicating Survival Mixed Regression model is better fit model compared to Cox model.

4. Discussion

From this study it is found that mothers of Chittagong division take their next child at lower interval compared to Barisal division. Similarly mothers living in Sylhet much more higher risk of occurring the next birth compared to the Barisal division. But mothers residing in Khulna and Rajshahi take their next child at a higher interval. Observing the p-value it may be concluded that the divisional effects on birth interval mentioned are significant [8, 14]. The mothers of rural area take their next child at a lower interval compared to the urban area (p value < 0.01). Muslim mothers have smaller birth interval compared to the non-Muslim mothers (p-value < 0.01). Similar results were found in the previous work [14]. This study shows that mothers having higher or secondary educated mothers have higher birth interval than the illiterate mothers. Similarly, primary educated mothers have larger spacing than non-educated mothers but much smaller than others. The results satisfies the previous findings that the length of birth interval increases with the increase level of mother's education [15]. In reality, mothers with low standard of living index have shorter birth interval. According to our analysis, we can say that mothers in both middle families and rich families take their next child at higher interval compared to poor class families [14]. This reflection were found in our study that wealth index is found to have significant effect (p-value < 0.01) on birth interval. Moreover, mothers who are exposure to media take their next child at higher interval compared to non-exposure. This findings were matched with the earlier study that media massages helped to create awareness to have small family norms [16]. This might suggests mother's working status doesn't have a potential impact on birth interval. However in Cox PH model it is found to be significant (p-value < 0.01) contradicts to the result found for SMR model (p-value > 0.01). This difference in the results might be due to the clustering effects. This result shows something different because in the previous studies it has been found that younger mothers tend to have shorter interval than older mothers [8]. But our study suggests that the mother's age at birth is found to have insignificant effect (p-value > 0.05) on birth interval. On the other hand, it is found to be significant in Cox PH model. Mother's having a daughter as the previous child take their next child at lower interval compared to male children (p-value < 0.01). So, these findings are found to have significant effect on birth interval which is similar to the previous studies [17]. It is observed that if a child is alive (p-value < 0.01) mothers would take their next child at higher interval compared to the mothers who have lost their child matched with the results [18].

5. Conclusion

The use of Survival Mixed Regression model and semi-parametric models (Cox PH model) allows us for the correction and comparative analysis for finding the effect of the variable of interest on the response (succeeding birth interval). Because of the elimination

Socioeconomic and Demographic Determinants of Birth Interval in Bangladesh

of the frailty term, Cox PH model lead to an erroneous conclusion which can be recovered through Survival Mixed Regression model. The Survival Mixed Regression model allows for the introduction of an observed frailty term that relaxes the independent assumption. Thus considering intra cluster effect in Survival Mixed Regression model the variables division, type of residence, mother's education, religion, wealth index, mother's access to media, sex of child, survival status of previous child are found to have significant effect and provide impact of this variable on birth interval. On the other hand, Cox PH model found to have different consequences such as for mother's working status variable and mother's age at birth variable. In demographic applications, all heterogeneity among data set is captured theoretically by observed covariates [19-20]. In this study, different individuals are found to have different amount of risk of experiencing the event and the reasons for this variability are captured by the relevant covariates.

Acknowledgement: We are very grateful to National Institute of Population Research and Training (NIPORT) for giving us the BDHS data access for our research work. Also, we are thankful to the reviewers for their comments to improve the presentation of the paper.

REFERENCES

1. G.Rodriguez, J.Hobcraft, J.Mc Donald and J.Trussell, *A Comparative Analysis of Determinants of Birth Intervals*, WFS, Netherland (1984).
2. D.Nath, D.Leonetti and M.Steele, Analysis of birth interval in a non-contraceptive Indian population: An evolutionary ecological approach, *J. Biosoc.*, 32(3) (2000) 343-354.
3. Y.Abdulkarimova, Frailty models for modeling heterogeneity. M.Sc Thesis, Mc Master University, Canada (2013).
4. J.J.Heckman, J.V.Hotz and Walker, New evidence of the timing and spacing on births, *The American Economic Review*, 75 (1985) 175-183.
5. J.Heckman and B.Singer, A method of minimizing the impact of distributional assumptions in economic models for duration data *Econometrical*, 52 (1984) 271-320.
6. J.Trussell and M.Rodriguez, Heterogeneity in demographic research: Convergent issues in Genetics and Demography. New York, Willey (1990).
7. S.Mahmood, B.Zainob and A.H.M.M.Latif, Frailty modeling for clustered survival data: An application to birth interval in Bangladesh, *J. Appld. Stat.*, 40(12) (2013) 2670-2680.
8. P.Hougard, Modeling heterogeneity in survival data, *Journal of Applied Probability*, 28(3) (1991) 695-701.
9. National Institute of Population Research and Training (NIPORT). Bangladesh Demographic and Health Survey 2014; Mitra and Associates, and ICF International Inc.: Dhaka, Bangladesh; Rockville, MD, USA, 2016
10. D.R.Cox, Regression model and life tables, *Journal of Royal Statistical Society Series*, B34 (1972) 187-220.
11. R.G.Gutierrez, Parametric frailty and shared frailty survival models, *The Stata Journal*, 2(1) (2002) 22-44.
12. J.R.Khan, W.Bari and A.H.M.M.Latif, Trend of determinants of birth interval dynamics in Bangladesh, *BMC Public Health*, 16(1) (2016) 934-945.

Sharlene Alauddin, Md. Akhtarul Islam and M. Ershadul Haque

13. S.Sharmin, B.Zainob, W.Bari and M.A.Islam, Exploring the recent changes in differential patterns of birth intervals in Bangladesh, *Social Science*, 709 (2012) 266-273.
14. N.Parr, Mass media promotion of family planning and use of modern contraception in Ghana. 24th IUSSP General Conference at Salvador, Bahia, Brazil (2001).
15. S.Mahmood and B.Zainob, An analysis of birth intervals in Bangladesh using product limit estimate and accelerated failure time regression model, *Dhaka University Journal of Science*, 59(1) (2011) 9-15.
16. H.T.Hoa, N.V.Toan, A.Johans, V.T.Hoa, B.Hoyer and L.A.Persson, Child spacing and two child policy in practice in rural Vietnam: Cross sectional survey, *British Medical Journal*, 313 (1996) 7065.
17. J.Trussell and T.Richards, Correcting for unmeasured heterogeneity in hazard models using Hwckman-Singer procedure, *Sociological Methodology*, 15 (1985) 242-276.
18. J.N.Hobcraft, J.W.McDonald and S.O.Rutstein, Demographic determinants of infant and early child mortality: a comparative analysis, *Population Studies*, 39(3) (1985) 363-385.