The Effect of Education on the Timing of Marriage and First Birth in Pakistan*

Lata Gangadharan and Pushkar Maitra[†]
February 2001

Abstract

A rapid rise in women's education levels, an increase in the age at marriage and an increase in the age at which they have their first child are key features of demographic transition in any country. Education is considered to be an essential component in this process because increases in educational attainment are likely to significantly affect both age at marriage and the duration between marriage and first birth - in particular increasing both the age at marriage and the time to first child. This paper uses individual level unit record data from Pakistan to examine the effect of education on the age at marriage and on the duration between marriage and first birth. We jointly estimate educational attainment, age at marriage and duration between marriage and first birth allowing for household level unobserved heterogeneity. Our estimation results show that ignoring correlation between the heterogeneity terms in the three main variables of interest results in inconsistent estimates. We find that educated women marry significantly later but education does not have any impact on the duration between marriage and first birth.

JEL Classification: J13, O10, C41, C24.

Key Words: Education, Age at Marriage, Age at First Birth, Asia, Pakistan.

1 Introduction

It is often argued that a rapid rise in educational attainment of women, an increase in the age at marriage and an increase in their age at which they have their first child are key features of demographic transition in any country. Caldwell (1980) postulates that the demographic transition that has happened in both developed countries over the last century

^{*}This research is funded by the Faculty Research Grant Scheme, Faculty of Economics and Commerce, University of Melbourne. We would like to thank Rajith Laxman for his excellent and painstaking research assistance. The paper has benefitted from conversations with Tim Fry and comments by Jack Caldwell, Sarmistha Pal and Sang-Hyop Lee. An earlier version of the paper was presented at the 2001 North American Econometric Society Winter Meetings held at New Orleans. The usual disclaimer applies.

[†]Lata Gangadharan, Department of Economics, University of Melbourne, VIC 3052, Australia. E-mail: gangadha@cupid.ecom.unimelb.edu.au. Pushkar Maitra, Department of Economics, Monash University, Clayton Campus, VIC 3168, Australia. E-mail: Pushkar.Maitra@BusEco.monash.edu.au

and in contemporary developing countries is primarily associated with an increase in educational attainment of women (also see Caldwell, Reddy and Caldwell (1983)). Education is regarded as being the prime catalyst in this process because increases in educational attainment are likely to significantly affect both age at marriage and the duration to first birth. It is suggested that increased education is likely to open up economic alternatives to getting married and bearing children. Increased education could imply that for women the utility of being single might exceed the utility of being married (Becker (1974)). Further women are likely to reduce their labour market participation after child bearing and hence the opportunity cost of child bearing could be higher for more educated women. So even after marriage the utility of delaying birth might exceed the utility of having a child immediately. A higher age at marriage is expected to be associated with a rise in the status of women, as they enter their reproductive life with greater maturity and capabilities. An increase in the age at marriage for women is also likely to reduce the age differential between the husband and the wife. This is likely to increase the bargaining power of the wife and reduce the power imbalance within the family. As with the age at marriage, age at first birth is also significant in the process of demographic transition because a higher age at first birth is typically associated with a lower lifetime fertility rate.

Over the last few decades' education levels have significantly increased in many countries as have the age at marriage and the duration between marriage and first birth. In this paper we use unit record data to examine whether Pakistan has undergone a demographic transition of this kind - increased educational attainment, increased age at marriage and increased duration between marriage and first birth. We use data from a national sample of women to examine the changes in educational attainment over the last five decades and the effect of education on the age at marriage and on the duration between marriage and first birth. We find that over the last fifty years educational attainment among women has increased significantly in Pakistan (the proportion of women having no education has decreased and the proportion of women having primary, secondary and higher secondary schooling has increased). However over the same time period, age at marriage does not show a significant increase nor does the duration between marriage and first birth. Actually the duration between marriage and first birth shows a marked decline. This paper makes a significant contribution to the analysis of demographic transition in Pakistan. While there are papers that examine issues relating to women's education and changes in fertility behaviour in Pakistan (Sathar (1984), Sathar and Kiani (1986, 1998), Sathar and Kazi (1989) and Sathar and Casterline (1998) to name a few), they are examined in isolation of each other. To the best of our knowledge, this is the first paper to examine the relationship between education, age at marriage and the duration between marriage and first birth in Pakistan and conduct a structural estimation that accounts for this inter-relationship.

In this paper we estimate a structural model that accounts for the interaction between the three main variables of interest - the level of education, the age at marriage and the duration between marriage and birth. The reason for joint estimation arises from the possible endogeneity of education in the age at marriage equation and the possible endogeneity of both education and the age at marriage in the duration to first birth equation. We find that education significantly reduces the hazard of marriage but does not have a significant direct effect on the duration between marriage and first birth. Further there are significant birth cohort effects on both the hazard of marriage and on the hazard of first birth.

The rest of the paper is organised as follows. Section 2 presents the econometric framework that we use in this paper. It sets out the structural model and discusses the estimation techniques used. Section 3 discusses the data and presents some selected descriptive statistics. Section 4 analyses the results and Section 5 concludes.

2 Econometric Framework

The estimation methodology used in this paper attempts to find the relationship between the education level attained by a woman, the age at which she gets married and the duration between marriage and first birth.

2.1 School Continuation

The first stage of the estimation process is to examine the highest level of education attained by a woman. This is estimated using a sequential probit model. Education outcomes are assumed to result from a series of sequential decisions. The first decision is whether to attend school at all. The second decision is conditional on the decision to attend school, decide whether to stop at grades 1 - 5 or continue beyond grade 5. The third decision is conditional on the decision to continue beyond primary school

The decision model is therefore formulated as a sequential probit model with an index function Z(s) for each decision as follows:

$$Z(s) = \alpha_0(s) + \alpha_1(s) X_1(s) + \zeta^e + \varepsilon^e(s)$$
(1)

Here $X_1(s)$ denotes a set of individual and household level decision variables that are exogenous, though they can potentially be endogenous. The heterogeneity component in the propensity to continue is given by ζ^e which is assumed to apply to all members of the same household and is constant across decisions. There is in addition a residual component given by $\varepsilon^e(s)$, which is specific to decision s and is independent from decision to decision. The heterogeneity and residual components are assumed to be normally and independently distributed so that $\zeta^e \sim N\left(0, \sigma_{\zeta}^2\right)$ and $\varepsilon^e(s) \sim IIDN(0,1)$ - IID across choices and households. A woman proceeds from level s to the next level s+1 if $Z_j(s)>0$. The sequential choices of individual j is determined by the corresponding ranges of the index function at each decision point:

$$\begin{array}{rcl} s & = & 1 \text{ if } Z\left(1\right) \leq 0 \\ s & = & 2 \text{ if } Z\left(1\right) > 0, Z\left(2\right) \leq 0 \\ & & \vdots \\ s & = & S-1 \text{ if } Z\left(1\right) > 0, Z\left(2\right) > 0, \dots, Z\left(S-2\right) > 0, Z\left(S-1\right) \leq 0 \\ s & = & S \text{ if } Z\left(1\right) > 0, Z\left(2\right) > 0, \dots, Z\left(S-2\right) > 0, Z\left(S-1\right) > 0 \\ & & 3 \end{array}$$

Some observations might be censored - some women might still be enrolled in school at the time of the survey. The probabilities of these censored events are given by the probability of not stopping at all the previous levels. So conditional on $X_1(s)$ and on ζ^e the continuation decisions are given by

$$\varepsilon^{e}(s) > \Theta(s, X_{1}(s)) + \zeta^{e}$$

for each decision point s where

$$\Theta\left(s, X_{1}\left(s\right)\right) = \alpha_{0}\left(s\right) + \alpha_{1}\left(s\right)X_{1}\left(s\right)$$

So the conditional probabilities of various completed and censored outcome levels are given by

$$L\left(\zeta^{e}\right) = \begin{cases} \prod_{k=1}^{s-1} \Phi\left(\Theta\left(k, X_{1}\left(k\right)\right) + \zeta^{e}\right) & \text{if censored at } s \\ \Phi\left(-\Theta\left(k, X_{1}\left(k\right)\right) + \zeta^{e}\right) \times \prod_{k=1}^{s-1} \Phi\left(\Theta\left(k, X_{1}\left(k\right)\right) + \zeta^{e}\right) & \text{if uncensored at } s < S \\ \prod_{k=1}^{S-1} \Phi\left(\Theta\left(k, X_{1}\left(k\right)\right) + \zeta^{e}\right) & \text{if uncensored at } S \end{cases}$$

$$(2)$$

The exogenous explanatory variables included are: the age of the woman, the square of the age of the woman, log of household income at the time of the survey, a dummy for marital status, dummies to indicate whether the woman's mother and father are literate, dummies for the primary occupation of the father (the reference category is that the father's occupation is missing), a dummy variable to capture whether the woman resides in a rural area and finally, three province dummies for residence in the North West Frontier Province, Sindh and Baluchistan. The reference category is that the household resides in Punjab.

2.2 Hazard Models for Age at Marriage and Duration between Marriage and First Birth

The age at marriage and the duration between marriage and first birth are both modelled as failure time processes represented by a log hazard of duration equations. The model is one of proportional hazards with covariates and heterogeneity shifting the baseline hazard. The log hazard equation is written as (see Brien and Lillard (1994)):

$$\ln h(t) = \xi_0 + \xi_1 T(t) + \xi_2 X + \varepsilon \tag{3}$$

where h(t) is the log hazard of an event at t, ξ_0 is a constant, T(t) is a vector of all duration dependent variables, X is a vector of other explanatory variables (some of which might be endogenous) and ε is a residual representing heterogeneity and we assume that $\varepsilon \sim N(0, \sigma_{\varepsilon}^2)$. T(t) is simply a spline in time since the person became at risk for the event (at risk of getting married or at risk of having a child). Let us denote the time at which a person enters the risk of an event by t_0 . The duration $t - t_0$ is divided into S discrete periods. Then the baseline log hazard function is defined as a spline or a piecewise linear

function and the log hazard of the event will have different slopes over the duration. So the baseline hazard is written as:

$$\xi_0 + \xi_1 T(t) = \xi_0 + \sum_{k=1}^{S+1} \xi_{1k} T_k(t)$$
(4)

In other words the baseline log hazard is the sum of the effects of the various sources of time dependence within the period of risk for an individual and the resulting log hazard equation is piecewise linear in time since the episode began.

2.2.1 Age at Marriage

The log hazard of marriage equation for a woman at time t is given by:

$$\ln h^{m}(t) = \beta_0 + \beta_1 AGE^*(t) + \beta_2 X_2 + \varepsilon^{m}$$
(5)

Here $AGE^*(t)$ measures duration dependence for age of the woman and X_2 denotes a vector of exogenous individual characteristics. $AGE^*(t)$, the duration dependence for age, measures the time varying risk of marriage from the time the woman is at the risk of getting married (age of 13). So for women that are married, $AGE^*(t)$ is Age at Marriage - 13, and for women that are not married, $AGE^*(t)$ equals the age of the women at the time of the survey - 13. The log hazard of marriage is therefore zero before the age of 13. The residual term ε^m represents unobserved heterogeneity, and is assumed to be normally distributed $\varepsilon^m \sim N(0, \sigma_m^2)$. The hazard equation for marriage is defined from the time the individual becomes at risk of marriage (we assume this to be age 13) until the time of marriage or until the time the process is censored by the survey date. The duration dependence for age of the woman enters as a spline with one node at 17 years of age. We tried using multiple nodes but the single node at 17 years of age fitted the data best. Further, the kernel density estimates for the age at marriage (see Figure 1) show that the peak of the distribution is attained around the age of 17. The exogenous variables that we include in this regression are four birth cohort splines: the woman was born before 1947, the woman was born between 1947 and 1960, the woman was born between 1960 and 1970 and finally the woman was born after 1970. We include two dummies to indicate whether the woman's parents are literate, three dummies to indicate the primary occupation of the father: Agriculture, Business or Other. The reference category for father's occupation is that father's occupation is missing. We tried to include the primary occupation of the mother but it led to problems with convergence of the maximum likelihood estimation procedure because there is very little variation in the primary occupation of the mother most did not work in the labour market. Finally we include two dummies for the highest level of education attained by the woman: the highest level of education attained by the woman is primary school and the highest level of education attained by the woman is secondary school or higher. The reference category is that the woman has no (formal) education.

Another variable of interest would be whether the woman was enrolled in school at the time of marriage as school enrolment could have an effect on the timing of marriage. However as this survey is not retrospective in nature, the respondents were not asked if they were enrolled in school at the time they got married. The only education data that we use is the stock of education, measured by the highest level of education attained by the woman. Flow variables like school enrolment status at the time of marriage or first birth are excluded from our analysis. The level of education attained could influence the age at marriage by increasing the opportunity cost of early marriage as more labour market opportunities are likely to be available for more educated women. At the same time, increasing the age at marriage, would free women to attain more education. Hence education levels are endogenous in the age at marriage equation and estimates could be biased unless this simultaneity is accounted for. The same argument holds for the hazard at first birth equation.

2.2.2 Duration Between Marriage and First Birth

Finally the log hazard equation that characterises the process leading to first birth by a woman is represented as follows:

$$\ln h^{c}(t) = \gamma_{0} + \gamma_{1}DURAT(t) + \gamma_{2}X_{3} + \varepsilon^{c}$$
(6)

Here DURAT (t) measures duration dependence for duration of marriage for the woman and X_3 denotes a vector of exogenous individual characteristics. The residual term ε^c is assumed to be normally distributed ($\varepsilon^c \sim N(0, \sigma_c^2)$) and represents unobserved heterogeneity. The hazard of first birth equation is defined from the time the individual becomes at risk of birth (as soon as the woman gets married) until the time of first birth or until the time the process is censored by the survey date. Notice that the birth hazard equation is defined only for women who are married. This is essentially due to the fact that in Pakistan, birth prior to marriage is not common and not socially accepted. In the estimation sample that we use in this paper, we ignore the few observed cases of child bearing prior to marriage and so the hazard of birth prior to marriage is effectively zero. Marriage duration enters as a spline with one node at 2 years following marriage. The kernel density estimates for the duration to first birth (see Figure 2) shows that the peak of the distribution is attained at around 2 years following marriage. Once again we tried using multiple nodes but this particular specification fitted the data best. The exogenous individual characteristics that we include in this regression are the following. First we include four birth cohort splines: the woman was born before 1947, the woman was born between 1947 and 1960, the woman was born between 1960 and 1970 and finally the woman was born after 1970. We include four age at marriage splines to account for the fact the age at marriage might affect the hazard of birth and further this effect could be non-linear. These splines are defined as married before the age of 15, married between the ages 15 and 20, married between the ages 20 and 30 and finally married after the age of 30. We include three dummies for the highest level of education attained by the husband: the highest level of education attained by the husband is primary school, the highest level of education attained by the husband is secondary school and the highest level of education attained by the husband is more than secondary school. As before we include two dummies to indicate the highest level of education attained by the woman. Finally we also control for household income at the time of survey, dummy for rural residence and three province dummies. Note that in the log hazard of the duration to first birth equation we do not include any of the parental characteristics, because in the South Asian context once a woman is married she moves to the husband's house and the woman's parents no longer have much influence in the decision making process.

In theory, we should also be including supply side factors in the above equations, for example, the availability of schools, hospitals and fertility clinics in the community could affect education levels and the time to first birth. However the community level characteristics capture supply side effects in 1991, when the survey was conducted, and not when the woman was being educated or getting married or having her first child. As data on these variables are not retrospective in nature, it is difficult to examine the impact of these supply side factors in this study. We include province dummies and these could help in part to capture differences in these kind of unobservable factors.

Table 1 summarises the information on the set of explanatory variables used in each of the three equations. Since equations (1), (5) and (6) are jointly estimated we have ensured that identification is attained.

The estimation is based on Full Information Maximum Likelihood following Lillard (1993) and Brien and Lillard (1994). Equations (1), (5) and (6) are jointly estimated and the heterogeneity components are assumed to be jointly normally distributed as follows:

$$\begin{bmatrix} \zeta^e \\ \varepsilon^m \\ \varepsilon^c \end{bmatrix} \sim N \begin{pmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\zeta}^2 & - & - \\ \rho_{\zeta m} \sigma_{\zeta} \sigma_m & \sigma_m^2 & - \\ \rho_{\zeta c} \sigma_{\zeta} \sigma_c & \rho_{mc} \sigma_m \sigma_c & \sigma_c^2 \end{bmatrix}$$
 (7)

The full marginal maximum likelihood for the model specified by equations (1), (5) and (6) is therefore given by:

$$L = \int_{\zeta^e} \int_{\varepsilon^m} \int_{\varepsilon^c} \frac{\phi\left(\frac{\zeta^e}{\sigma_\zeta}, \frac{\varepsilon^m}{\sigma_m}, \frac{\varepsilon^c}{\sigma_c} \middle| \rho_{\zeta m}, \rho_{\zeta c}, \rho_{mc}\right)}{\sigma_\zeta \sigma_m \sigma_c} \tag{8}$$

Additionally there is unobserved heterogeneity arising from the fact that several women might belong to the same household and there might be some household level characteristics (unobserved to the researcher) that affect the three decision variables. There is also a potential unobserved heterogeneity arising from sibling effects - for example multiple women in the sample could presumably be born of the same mother and there might be common mother level characteristics, again unobserved to the researcher, across these women. In our estimation we were able to account for the household level unobserved heterogeneity but not the sibling level unobserved heterogeneity because of unavailability of data. Estimation is based on maximisation of the marginal likelihood and is conducted using the Applied Maximum Likelihood (AML) Software Package (Lillard and Panis (2000)).

3 Data and Descriptive Statistics

The data set used is from the Pakistan Integrated Household Survey conducted in 1991 (PIHS). This survey was conducted jointly by the Federal Bureau of Statistics, the Government of Pakistan and the World Bank as a part of the Living Standard Measurement Study (LSMS) household surveys in a number of developing countries. The PIHS teams visited 4800 households residing in rural and urban communities. The nation-wide survey gathered individual and household level data using a multi-purpose questionnaire.

In this paper we use data on women more than 13 years of age. The sample consists of 6928 women from 4069 households. It is clear therefore that there are multiple women from the same household. In particular the data shows that there are 2296 (56.43%) households with one woman in the sample, 1050 (25.80%) households with two women in the sample, 471 (11.58) households with three women in the sample and the remaining (8.19%) households with more than 3 women in the sample. Ignoring household level unobserved heterogeneity could therefore lead to inconsistent estimates. Selected descriptive statistics are presented in Table 2. The average age of the women in the sample is 27.58 years while the average age of the husbands in the sample is 35.73 years. The average years of schooling for women in the sample is only 1.77 years and actually more than 75% of the women have no (formal) education at all. 43.56% of the women in the sample have husbands with no education, only 16.82% of the women have literate fathers and only 3.88% of the women have literate mothers. Finally of the 6928 women in the sample, 4546 are married and 3999 have had at least one child. The average age at marriage is 17.51 years and the average duration to first birth is 3.21 years (giving an average age at first birth of approximately 20.5 years). The average duration to first birth is only computed for women that are married because in Pakistan being married is a pre-condition to having children. Figures 1 - 3 present the kernel density estimates of the age at marriage, the age at first birth and the duration between marriage and first birth respectively. Notice that the peak of the distribution of the age at marriage is attained around the age of 17, while the peak of the distribution of the duration between marriage and first birth is attained around 2 years following marriage.

Figures 4 and 5 present selected evidence on educational attainment of women in the sample, classified by birth cohorts. The average years of education has increased from 0.77 years for women born prior to 1945 to 2.66 years for women born during the years 1965 - 1970, but there has been a dip in the number of years of school attended for women born after 1970 (see Figure 4). Figure 5 presents the proportion of women in each of the three educational groups: primary school educated, secondary school educated and higher secondary/college educated. We find that the proportion of women with primary schooling reaches a peak for women born during the period 1965 - 1970 and then falls. Starting from very low levels of female education - approximately 10% of women born prior to 1945 have any kind of formal education, this figure rises to more than 32% for women born between 1965 and 1970 and the figure then goes down to around 24% for women born after 1970.

Turning to the average age at marriage (see Figure 6), we see that the average age at marriage has remained roughly the same for women born during the period 1945 -

1970 but falls significantly for women born after 1970, though it must be noted that more than 77% of the women born after 1970 were not married at the time of the survey. We also computed the average age at marriage, by marriage cohorts. These results are not presented but are available on request. The descriptive statistics show that the average age at marriage has increased from 15.15 years for women married prior to 1960 to 18.50 years for women married after 1985. So while the average age at marriage has increased over the last forty years, the increase is significantly smaller than those presented by Sathar and Kiani (1986) who find that the average age at marriage has increased from approximately 16 years for women married before 1960 to 22 years for women married after 1980. Both the average age at first birth and the duration between marriage and first birth has fallen significantly over the sample period. For example while the average duration to first birth (after marriage) was 5.71 years for women born after 1945 it fell to around 2.05 years for women born after 1970. This fall in duration between marriage and first birth is similar to those obtained by Donaldson and Nichols (1978) and Rindfuss and Morgan (1983) who argue that it is the increase in the age at marriage that has resulted in a significant decrease in the duration between marriage and first birth. Feng and Quanhe (1996) find that in China age at marriage has gone up significantly and the duration between marriage and first birth has gone down significantly. They attribute this to the changing sexual behaviour among Chinese couples. Brien and Lillard (1994) find that in Malaysia while the age at marriage and the age at first conception have both increased, the former has increased at a faster rate and the average time to first conception has fallen.

4 Results

The full set of results for the 55 jointly estimated parameters for equations (1), (5) and (6) are presented in Tables 3 - 6. The complete model will be referred to as Model III. For comparison two other models are also presented, where the correlation across the equations is restricted to be zero. Additionally in one of the models (referred to as Model I) we do not control for unobserved heterogeneity, which we however do in the second model (Model II). Table 3 presents the standard errors and correlations for the heterogeneity components. Table 4 presents the sequential probit estimates for school continuation the highest level of schooling completed, Table 5 presents the coefficient estimates for the log hazard of marriage equation and finally Table 6 presents the coefficient estimates for the log hazard of birth equation. A positive coefficient in Table 5 increases the hazard of marriage and reduces the age at marriage while a negative coefficient decreases the hazard of marriage and increases the age at marriage. Correspondingly, a positive coefficient in Table 6 increases the hazard of birth and reduces the duration between marriage and first birth while a negative coefficient reduces the hazard of birth and increases the duration between marriage and first birth.

The problem of endogeneity arises when an explanatory variable in one of the processes is the outcome of a potentially related process. In such a situation if the heterogeneity terms are correlated, then the estimates are likely to be biased particularly if this correlation is ignored. Failure to incorporate the endogeneity and the corresponding correlation among

the different processes also results in inconsistent standard errors, particularly when the correlations are significant. Table 3 presents the estimated standard deviations and the correlations of the heterogeneity components. These estimates show that the household level unobserved heterogeneity $(\sigma_{\zeta}, \sigma_{m}, \sigma_{c})$ are large and statistically significant at 1% level in each of the three processes. Further, only one of the three correlations between the processes is statistically significant and that is the correlation between the education and marriage equations $(\rho_{\zeta m})$.

Why should the three heterogeneity terms be correlated - i.e., why do we expect $\rho_{\zeta m} \neq 0$, $\rho_{\zeta c} \neq 0$ and $\rho_{mc} \neq 0$? This is because more educated women are likely to have more job opportunities and hence the opportunity cost of marriage and child bearing could be higher for more educated women. Therefore some of the unobservable characteristics that affect the school continuation decisions could also affect the hazard of marriage and the hazard of bearing a child after marriage and vice versa.

How much of a difference does ignoring the potential bias cause? Notice that the highest level of completed education of the woman has no effect on the duration between marriage and first birth, irrespective of which model we use (Table 6), while the highest level of completed education has a significant positive effect on the age at marriage (Table 5). Interestingly the effect of the woman's education on the marriage hazard is the strongest in Model III. For example, the estimated coefficients indicate that for the complete model (Model III), relative to a woman with no education, a woman with primary schooling as the highest level of education attained has a 29% lower hazard of marriage, while a woman with secondary schooling or more has a 77% lower hazard of marriage. The effects are underestimated in both Model I and Model II. A similar underestimation of effect is evident in the school continuation equation (Table 4). For example women who are not married have a significantly higher probability of continuing to remain in school beyond the secondary school level. This effect is underestimated in Model I.

Let us now turn to a detailed examination of the results. Starting with the education equation (Table 4), we find that older women are more likely to have attended (any) school, completed primary school and completed secondary school, though there is a significant non-linearity in the age effect on the school continuation probabilities. Married women on the other hand have a significantly lower probability of attending some school and completing primary and secondary school. Women with literate parents have a significantly higher probability of continuing in school while a woman whose father's primary occupation is in agriculture has a significantly lower probability of ever attending school and a significantly lower conditional probability of completing primary school. Relative to women living in urban areas, women living in rural areas have a significantly higher probability of having no schooling as do women living in Sindh, NWFP and Baluchistan (relative to women living in the Punjab province). Women belonging to richer households have a higher probability of continuing in school.

The estimates for the log hazard at marriage equation are presented in Table 5. The hazard of marriage rises significantly in the first 4 years of eligibility of marriage (remember that we assume that a woman becomes eligible for marriage at the age of 13) and falls

¹This is computed by $1 - e^{-\beta}$ where β is the estimated coefficient.

thereafter. The slope is however considerably flatter after the age of 17. The hazard of marriage is significantly lower for women with literate parents. In fact the estimated coefficients show that the hazard of marriage is lower by 23% for women with literate mothers and is lower by 11% for women with literate fathers. Both of these essentially translate to a higher age at marriage for women with literate parents. There are significant birth cohort effects. The configuration of the sign of the birth cohort effects is somewhat curious. The hazard of marriage is higher for women born prior to 1947 and for women born after 1960 (the estimated coefficients are positive and significant) while the hazard of marriage is lower for women born between 1947 and 1960. This configuration of signs implies that the age at marriage was significantly lower for women born before 1947 and after 1960 but was significantly higher for women born between 1947 and 1960. The estimated coefficients match the trends presented in Figure 2 where we find a sharp increase in the average age of marriage for women born between 1945 and 1960 and the average age falls thereafter. One reason for this significant increase in the age of marriage for women born between 1947 and 1960 could be the low men to women ratio in the population - a possible consequence of the partition of India in 1947 and the riots that followed. However one cannot elaborate on this without more vital statistics evidence, which is unfortunately not available for Pakistan during the period under consideration. This issue is explored in greater detail by Sathar and Kiani (1998) who argue that the population structure in Pakistan is quite bottom-heavy. Given that husbands are likely to be older than wives, this leads to a population squeeze, with fewer men per women at every marriage cohort and this effects the age at marriage. Finally, as argued already, there are significant effects of continuing in school on the hazard of marriage. The age of marriage is significantly higher for women with primary schooling and even higher for women with secondary schooling or more. Education increases the opportunity cost of earlier marriage as more labour market opportunities are likely to be available for more educated women.

Turning to the duration between marriage and first birth (Table 6) we find that the hazard of birth is increasing in the first two years following marriage and falls thereafter. These results are consistent with the kernel density estimates of the duration between marriage and first birth (Figure 6), which shows that the peak of the distribution is attained around two years following marriage. There are significant age at marriage effects on the hazard of birth and hence on the duration between marriage and first birth. configuration of signs is however somewhat surprising. The hazard of birth is significantly higher for women who marry before the age of 15 and between the ages 15 and 20 but is significantly lower for women who marry after the age of 30. Therefore the time to first birth (following marriage) is lower for women marrying before the age of 20 and is higher for women marrying after the age of 30. There are again significant birth cohort effects. The estimated coefficients show that the duration between marriage and first birth is significantly lower for women born prior to 1970. Finally the highest level of education attained by the woman herself and the highest level of education attained by the husband have little or no effects on the duration between marriage and first birth. Why should these two results be of interest? It could be argued that the characteristics of a husband and wife are endogenous in a marriage in that men with certain characteristics choose women with certain characteristics to be their wives. It could be, for example, the case that men with more education marry more educated women and hence marry women who are older and these women have a lower duration between marriage and birth. However we do not find any evidence of this kind of behaviour and generally the level of education of the husband has very little effect on the hazard of birth. What about the effect of education of the woman on the hazard of birth? As education levels increase so do the employment opportunities. The opportunity cost of early birth therefore increases because in many cases women exit the labour force at the time of child bearing. Their employment careers experience a break, which results in lower lifetime earnings. An increase in educational attainment therefore should be associated with an increase in the duration between marriage and first birth. In Pakistan we do not find such a relationship. This could be explained in part by the fact that even though the education level of women might have increased, women are not putting this increased education to use in the labour market. This is reflected in significantly lower female labour market participation rates in Pakistan (approximately 20% in 1980) relative to the average for the developing countries (around 40% in 1980) and also relative to the average for the South Asian countries (around 34% in 1980).

5 Conclusion

In this paper we estimate a structural model of educational attainment, age at marriage and age at first birth among women in Pakistan. The reason for joint estimation arises from the potential endogeneity of education in the age at marriage equation and the endogeneity of both education and age at marriage in the duration to first child equation. This leads to correlation between the errors in the three principal equations and if the correlation is not accounted for, that could lead to inconsistent and biased estimates. The estimation methodology used in this paper allows us to account for unobserved household level heterogeneity, which we find to be always significant though the correlation between the errors is significant only in one case - between the highest level of education attained and the age at marriage.

Our results show that the highest level of education attained by a woman has a significant effect on the age at marriage but has no effect on the duration between marriage and first birth. This could be because of the fact that women, even if they are highly educated, do not use this education in the labour market to earn a higher premium on their wages. Thus the opportunity cost of child bearing is possibly not very high in Pakistan.

We also find that the hazard of marriage reaches a peak around the age of 17 - the hazard increases in the first four years following the entry of a woman into the risk of being married category. The duration between marriage and first birth reaches a peak at two years following marriage. Additionally we find significant birth cohort effects on both age at marriage and the duration between marriage and first birth. Finally parental education has a significant effect on the highest level of education attained by a woman and on the age at marriage, but the education level of the husband has no effect on the duration between marriage and first birth.

Bibliography

- 1. Anderson, K. H., Hill, M. A. and Butler, J. S. (1987), 'Age at Marriage in Malaysia: A Hazard Model on Marriage Timing', Journal of Development Economics, 26(2), pp 223 234.
- 2. Becker, G. S. (1974), 'A Theory of Marriage' in Economics of the Family: Marriage, Children and Human Capital, T. W. Schultz (Ed), pp 299 344, Chicago, University of Chicago Press.
- 3. Caldwell, J. C., Reddy, P. H. and Caldwell, P. (1982), 'The Causes of Demographic Change in Rural South India: A Micro Approach', Population and Development Review, 8(4), pp 689-727.
- 4. Caldwell, J. C. (1980), 'Mass Education as a Determinant of the Timing of Fertility Decline', Population and Development Review, 62(2), pp 225-255.
- 5. Brien, M. J. and Lillard, L, A. (1994), 'Education, Marriage, and First Conception in Malaysia', Journal of Human Resources, 29(4), pp 1166 1204.
- 6. Brien, M. J., Lillard, L. A. and Waite, L. J. (1999), 'Interrelated Family-Building Behaviors: Cohabitation, Marriage and Nonmarital Conception', Demography, 36(4), pp 535 551.
- 7. Donaldson, P. J. and Nichols, D. J. (1978), 'The Changing Tempo of Fertility in Korea', Population Studies, 32(2), pp 231 250.
- 8. Feng, W. and Quanhe, Y. (1996), 'Age at Marriage and the First Birth Interval: The Emerging Change in Sexual Behavior Among Young Couples in China', Population and Development Review, 22(2), pp 299 320.
- 9. Gangadharan, L. and Maitra, P. (2001), 'Two Aspects of Fertility Behaviour in South Africa', Economic Development and Cultural Change, July. Forthcoming.
- 10. Greene, W. H. (1990), Econometric Analysis, Macmillan Publishing Company.
- 11. Lillard, L. A. and Panis, C. W. A. (2000), Applied Maximum Likelihood User's Guide and Reference Manual, Econware, Inc.
- 12. Lillard, L. A. (1993), 'Simultaneous Equations for Hazards: Marriage Duration and Fertility Timing', Journal of Econometrics, 56(1-2). pp 189 217.
- 13. Lee, S-H. (2000), 'Demand for Immunization and Child Survival: Evidence from Rural India', Mimeo, University of Hawaii.
- 14. Raut, L. K. (1996), 'Old-Age Security and Gender Preference Hypothesis: A Duration Analysis of Malaysian Family Life Survey Data', Journal of Quantitative Economics, 12(1), pp 81 104.

- 15. Rindfuss, R. R. and Morgan, S. P. (1983), 'Marriage, Sex and the First Birth Interval: The Quiet Revolution in Asia', Population and Development Review, 9(2), pp 259 278.
- 16. Sathar, Z. A. (1984), 'Does Female Education Affect Fertility Behaviour in Pakistan?', Pakistan Development Review, 23(4). pp 573 90.
- 17. Sathar, Z. A. and Casterline, J. B., (1998), 'The Onset of Fertility Transition in Pakistan' Population and Development Review, 24(4), pp 773 796.
- 18. Sathar, Z. A. and Kiani, M. F. (1998), 'Some Consequences of Rising Age at Marriage in Pakistan', Pakistan Development Review, 37(4), pp 541 54.
- 19. Sathar, Z. A. and Kazi, S. (1989), 'Female Employment and Fertility: Further Investigation of an Ambivalent Association', Pakistan Development Review, 28(3), pp 175 93.
- 20. Sathar, Z. A. and Kiani, M F. (1986), 'Delayed Marriages in Pakistan', Pakistan Development Review, 24(4).
- 21. Subbarao, K. and Raney, L. (1995), 'Social Gains from Female Education: A Cross-National Study', Economic Development and Cultural Change, 44(1), pp 105 128.

Table 1: Explanatory Variables Used

| Variable | Education | Marriage | Birth |
|---|-----------|----------|-------|
| Age of Woman | X | | |
| Square of Age of Woman | X | | |
| Marital Status | X | | |
| Mother Literate | X | X | |
| Father Literate | X | X | |
| Rural Residence | X | | X |
| Province of Residence | X | | X |
| Household Income | X | | X |
| Duration of Risk of Marriage (Age Spline) | | X | |
| Father's Primary Occupation | X | X | |
| Birth Cohort Spline | | X | X |
| Highest Education Attained by Woman | | X | X |
| Highest Education Attained by Husband | | | X |
| Marriage Duration Spline | | | X |
| Age at Marriage Spline | | | X |

Table 2: Selected Descriptive Statistics

| Variable | Number of Observations | Mean |
|------------------------------|-------------------------|----------|
| Household Level Variables | | |
| Rural Residence | 4069 | 1.495945 |
| Household Income | 4069 | 56909.43 |
| Household Size | 4069 | 7.805849 |
| Individual Level Variables | | |
| Years of Schooling | 6928 | 1.778435 |
| Proportion of Women with: | | |
| No Education | 6928 | 0.76198 |
| Primary Schooling | 6928 | 0.23802 |
| Secondary Schooling | 6928 | 0.132939 |
| Higher Secondary Schooling | 6928 | 0.039405 |
| Highest Level of Education A | ttained by the Husband: | |
| Primary Schooling | 6928 | 0.182015 |
| Secondary Schooling | 6928 | 0.280456 |
| Higher Secondary Schooling | 6928 | 0.100895 |
| No Schooling | 6928 | 0.435624 |
| Age of Husband | 6928 | 35.7295 |
| Age of Woman | 6928 | 27.58054 |
| Not Married | 6928 | 0.343822 |
| Age at Marriage* | 4546 | 17.50748 |
| Not conceived* | 4546 | 0.120326 |
| Age at Birth** | 3999 | 20.50413 |
| Duration to Birth** | 3999 | 3.205972 |
| Father's Occupation: | | |
| Agriculture | 6928 | 0.272373 |
| Business | 6928 | 0.129042 |
| Other | 6928 | 0.287818 |
| Missing | 6928 | 0.310768 |
| Father Literate | 6928 | 0.168158 |
| Mother Literate | 6928 | 0.038828 |

^{*:} Average Computed for those women that are married at the time of the survey

^{**:} Average Computed for those women that have at least one child at the time of the survey

Table 3: Heterogeneity Structure

| Variable | Model I | Model II | Model III |
|--|---------|-----------|-----------|
| $\sigma_{\zeta}(\text{School Continuation})$ | - | 0.5785*** | 0.5835*** |
| | | (0.0407) | (0.0412) |
| $\sigma_m(\text{Age at Marriage})$ | - | 0.8583*** | 0.8129*** |
| | | | (0.0408) |
| σ_c (Duration between Marriage and First Birth) | - | 0.4771*** | 0.4763*** |
| | | (0.0559) | (0.0571) |
| $ ho_{\zeta_m}$ | - | - | 0.2082* |
| | | | (0.1259) |
| $ ho_{\zeta c}$ | - | - | 0.0217 |
| | | | (0.077) |
| $ ho_{mc}$ | - | - | -0.0586 |
| | | | (0.1124) |
| Notes: | | | |
| Asymptotic Standard Errors in Parenthesis | | | |
| Significance: *: 10%; **: 5%; ***: 1% | | | |

Table 4: Sequential Probit Estimates of School Continuation Attending Primary School

| Variable | Model I | Model II | |
|--------------------------|------------|------------------------|------------|
| Constant | -1.8775*** | -2.2163*** | -2.3126*** |
| | (0.1863) | (0.249) | (0.2514) |
| Age of Woman | 0.0255*** | 0.0216*** | 0.0285*** |
| | (0.0051) | (0.0058) | (0.0059) |
| Square of Age of Woman | -0.0005*** | -0.0004*** | -0.0005*** |
| | (0.0000) | (0.0000) | (0.0000) |
| Married | -0.2498*** | -0.2094*** | -0.2196*** |
| | (0.0632) | (0.0758) | (0.0774) |
| Mother Literate | 0.7423*** | 0.8570*** | 0.8579*** |
| | (0.0832) | (0.0964) | (0.0995) |
| Father Literate | 0.8675*** | 0.9504*** | 0.9509*** |
| | (0.052) | (0.0618) | (0.0626) |
| Father's Education | | | |
| Agriculture | | -0.5394*** | |
| | (0.0717) | (0.087) -0.2394*** | (0.0884) |
| Business | -0.1927*** | -0.2394*** | -0.2410*** |
| | (0.0743) | (0.0906) | (0.0924) |
| Other | -0.2515*** | -0.2918*** | -0.2899*** |
| | (0.0645) | (0.0799) | (0.0816) |
| | | | |
| Rural Residence | -0.5302*** | -0.6423*** | -0.6434*** |
| | (0.0385) | (0.0533) | (0.054) |
| Residence in Sindh | -0.2308*** | | |
| | (0.0408) | (0.0558) -0.6728*** | (0.0568) |
| Residence in NWFP | -0.5907*** | -0.6728*** | -0.6731*** |
| | (0.0538) | (0.0723) | (0.0729) |
| Residence in Baluchistan | -0.8179*** | | |
| | | (0.1221) | |
| Log Household Income | | 0.1687*** | |
| | (0.0153) | (0.0214) | (0.0217) |
| | | | |

Table 4 (Continued): Sequential Probit Estimates of School Continuation Completing Primary School

| Variable . | | Model II | |
|--------------------------|------------|------------------------|------------|
| Constant | | -5.0760*** | |
| | (0.4612) | (0.5474) | (0.5515) |
| Age of Woman | 0.2122*** | (0.5474) $0.2204***$ | 0.2210*** |
| | (0.0265) | (0.0302) -0.0034*** | (0.0306) |
| Square of Age of Woman | -0.0032*** | -0.0034*** | -0.0034*** |
| | | (0.0005) | (0.0005) |
| Married | -0.7868*** | -0.8864*** | -0.8836*** |
| | (0.1304) | (0.1483) | (0.1509) |
| Mother Literate | 0.7200*** | 0.9484*** | 0.9521*** |
| | (0.1264) | (0.1473) $0.9175****$ | (0.1477) |
| Father Literate | 0.6910*** | 0.9175*** | 0.9219*** |
| | (0.0915) | (0.1044) | (0.1058) |
| Father's Occupation | | | |
| Agriculture | | -0.4848*** | |
| | (0.1477) | (0.1662) | (0.1689) |
| Business | -0.1933 | -0.2344 | -0.2462 |
| | (0.1423) | (0.1609) | (0.1633) |
| Other | -0.3120** | -0.3648** | -0.3725** |
| | | (0.1421) | |
| | | | |
| Rural Residence | -0.7418*** | -1.0266*** | -1.0312*** |
| | (0.0756) | | |
| Residence in Sindh | | 0.1091 | |
| | | (0.0921) | |
| Residence in NWFP | 0.1453 | 0.0228 | 0.0262 |
| | (0.1164) | (0.1376) | (0.1388) |
| Residence in Baluchistan | -0.1171 | -0.4022* | -0.4104* |
| | (0.1902) | (0.2315) | (0.2339) |
| Log Household Income | 0.1432*** | 0.2113*** | 0.2152*** |
| | | (0.0387) | |

Table 4 (Continued): Sequential Probit Estimates of School Continuation Completing Secondary School

| Variable | Model I | Model II | Model III |
|--------------------------|--------------|----------------------|------------|
| Constant | -2.9763*** | -4.5131*** | -4.6102*** |
| | (0.8422) | (0.9596) | (0.9709) |
| Age of Woman | 0.0992** | | 0.1223** |
| | (0.0487) | (0.0539) | (0.0545) |
| Square of Age of Woman | -0.0015* | -0.0017** | -0.0018** |
| | (0.0008) | (0.0008) | (0.0008) |
| Married | -0.6709*** | -0.8400*** | -0.8418*** |
| | (0.1728) | (0.1977) | (0.2015) |
| Mother Literate | 0.4188*** | 0.6384*** | 0.6482*** |
| | (0.1447) | (0.1607) | (0.1658) |
| Father Literate | 0.4312** | (0.1607) $0.6896***$ | 0.6852*** |
| | | (0.2014) | |
| Father's Occupation | | | |
| Agriculture | -0.3235 | -0.3708 | -0.3767 |
| | (0.3021) | (0.3222) | (0.3237) |
| Business | -0.2778 | -0.337 | -0.346 |
| | (0.2193) | (0.2547) | (0.2569) |
| Other | -0.1502 | -0.2252 | -0.2291 |
| | (0.2188) | (0.2423) | (0.2448) |
| Rural Residence | -0.2946 | -0.6916*** | -0.6981*** |
| | (0.1929) | (0.2194) | (0.2208) |
| Residence in Sindh | 0.1763 | 0.1428 | , |
| | (0.111) | (0.1291) | (0.1297) |
| Residence in NWFP | -0.1042 | | |
| | (0.1689) | | |
| Residence in Baluchistan | $0.0167^{'}$ | -0.3042 | , |
| | (0.3748) | (0.4664) | (0.472) |
| Log Household Income | 0.1351*** | | |
| | (0.0458) | (0.0525) | (0.0528) |
| N.I | | | |

Table 5: Log Hazard of Marriage

| Variable | Model I | Model II | Model III |
|----------------------------|-------------|-------------|------------|
| Constant | -21.5968*** | -27.8801*** | -7.3918*** |
| | (0.3728) | (0.6348) | (0.6485) |
| Age Spline | | | |
| Age 13 - 17 | 7.1716*** | 9.1291*** | 9.0499*** |
| | (0.116) | (0.1981) | (0.2024) |
| Age > 17 | -2.3715*** | -0.3797 | -0.5237* |
| | (0.2858) | (0.2706) | (0.2854) |
| N | 0.000 | 0.00==44 | 0.0000 |
| Mother Literate | -0.2095** | -0.2855** | -0.2775** |
| | (0.0931) | (0.1176) | (0.1235) |
| Father Literate | -0.1008** | -0.1576*** | -0.1298** |
| | (0.0409) | (0.056) | (0.0571) |
| Father's Occupation | | | |
| Agriculture | 0.6380*** | 0.7669*** | 0.7533*** |
| | (0.0808) | (0.1026) | (0.1012) |
| Business | 0.7202*** | 0.8736*** | 0.8651*** |
| | (0.0861) | (0.1113) | (0.1103) |
| Other | 0.7263*** | 0.8859*** | 0.8721*** |
| | (0.0825) | (0.1046) | (0.1028) |
| Birth Cohort Splines | | | |
| Prior to 1947 | 0.0424*** | 0.0630*** | 0.0564*** |
| | (0.0042) | (0.0055) | (0.0055) |
| 1947 - 1960 | -0.0164*** | -0.0271*** | -0.0148*** |
| | (0.0038) | (0.0051) | (0.0052) |
| 1960 - 1970 | 0.0764*** | 0.0750*** | 0.0368*** |
| | (0.0055) | (0.007) | (0.0086) |
| After 1970 | 1.0961*** | 1.2786*** | 0.3003*** |
| | (0.4086) | (0.3458) | (0.0401) |
| Highest Education Attained | • | , | |
| Primary School | -0.1845*** | -0.2742*** | -0.2874*** |
| | (0.056) | (0.076) | (0.0826) |
| More than Primary School | -0.5139*** | -0.7596*** | -0.7677*** |
| | (0.0666) | (0.0885) | (0.1024) |
| Notes: | | | |

Table 6: Log Hazard of First Birth

| Variable | | Model II | Model III |
|---|------------|-------------------|-------------|
| Constant | -8.8386*** | -10.5591*** | -10.5706*** |
| | (0.6862) | (0.8592) | (0.8764) |
| Marriage Duration Splines | | | |
| $\leq 2 \text{ years}$ | 0.5528*** | 0.7821*** | 0.7821*** |
| | (0.0317) | (0.0559) | (0.0565) |
| >2 years | -0.1532 | -0.3264** | -0.3247** |
| | (0.1206) | (0.1281) | (0.1291) |
| D 1D 11 | 0.100=*** | 0 1 - 1 1 4 4 4 4 | 0.1000444 |
| Rural Residence | -0.1287*** | -0.1514*** | -0.1666*** |
| | (0.034) | \ / | (0.0431) |
| Residence in Sindh | -0.2121*** | | -0.2243*** |
| | (0.0377) | (0.0469) | (0.0474) |
| Residence in NWFP | 0.1815*** | 0.2199*** | 0.2086*** |
| | (0.0432) | | (0.0558) |
| Residence in Baluchistan | -0.2884*** | -0.3103*** | -0.3277*** |
| | (0.0716) | | (0.0883) |
| Log of Household Income | 0.0666*** | 0.0780*** | 0.0826*** |
| | (0.0143) | (0.0175) | (0.018) |
| Age at Marriage Splines | | | |
| ≤ 15 | 0.1632*** | 0.2022*** | 0.1986*** |
| | (0.0332) | (0.0379) | (0.0385) |
| 15 - 20 | 0.0420*** | 0.0497*** | 0.0441*** |
| | (0.0099) | (0.0118) | (0.0154) |
| 20 - 30 | -0.0046 | 0.006 | 0.0028 |
| | (0.0089) | (0.0126) | (0.0149) |
| ≥ 30 | -0.2936*** | -0.3753*** | -0.3773*** |
| | (0.1128) | (0.1168) | (0.1184) |
| Notes: | | | |
| Asymptotic Standard Errors in Parenthesis | | | |

Table 6 (Continued): Log Hazard of First Birth

| Variable | Model I | Model II | Model III | |
|---|-----------|-----------|-----------|--|
| Highest Education of Husband | | | | |
| Some Primary School | 0.0839* | 0.0908* | 0.1086* | |
| | (0.0448) | (0.0526) | (0.0561) | |
| Middle School | 0.0561 | 0.0673 | 0.0718 | |
| | (0.0403) | (0.0486) | (0.0526) | |
| Above Secondary School | 0.0909 | 0.0889 | 0.1332 | |
| | (0.061) | (0.0725) | (0.0811) | |
| Year of Birth Splines | | | | |
| Prior to 1947 | 0.1067*** | 0.1250*** | 0.1260*** | |
| | (0.0108) | (0.013) | (0.0132) | |
| 1947 - 1960 | 0.0165*** | 0.0185*** | 0.0190*** | |
| | (0.0047) | (0.0054) | (0.0055) | |
| 1960 - 1970 | 0.0373*** | 0.0395*** | 0.0393*** | |
| | (0.0067) | (0.0078) | (0.008) | |
| After 1970 | -0.0161 | -0.0189 | -0.0191 | |
| | (0.0328) | (0.0362) | (0.0369) | |
| Highest Education Attained | | | | |
| Primary School | 0.0112 | 0.0049 | -0.0541 | |
| | (0.0565) | (0.0676) | (0.0772) | |
| More than Primary School | 0.0348 | 0.0518 | -0.0442 | |
| | (0.0523) | (0.0663) | (0.093) | |
| Notes: | | | | |
| Asymptotic Standard Errors in Parenthesis | | | | |
| C'. 'C * 1007 ** F07 *** 107 | | | | |

Significance: *: 10%; **: 5%; ***: 1%