Sex Selective Abortions, Fertility, and Birth Spacing*

Claus C Pörtner
Department of Economics
Albers School of Business and Economics
Seattle University, P.O. Box 222000
Seattle, WA 98122
work@clausportner.com
www.clausportner.com
&
Center for Studies in Demography and Ecology
University of Washington

October 2014

*I am grateful to Andrew Foster and Darryl Holman for discussions about the method. I owe thanks to Shelly Lundberg, Daniel Rees, David Ribar, Hendrik Wolff, seminar participants at University of Copenhagen, University of Michigan, University of Washington, University of Århus, the Fourth Annual Conference on Population, Reproductive Health, and Economic Development, and the Economic Demography Workshop for helpful suggestions and comments. I would also like to thank Nalina Varanasi for research assistance. Support from Danish Trust Fund resources in the World Bank and the University of Washington Royalty Research Fund is gratefully acknowledged. Partial support for this research also came from a Eunice Kennedy Shriver National Institute of Child Health and Human Development research infrastructure grant, 5R24HD042828, to the Center for Studies in Demography and Ecology at the University of Washington. Prior versions of this paper were presented under the title "The Determinants of Sex Selective Abortions." The views and findings expressed here are those of the author and should not be attributed to the World Bank or any of its member countries.

Abstract

Previous research on sex selective abortions has ignored the interactions between fertility, birth spacing, and sex selection, despite both fertility and birth spacing being important considerations for parents when deciding on the use of sex selection. This paper presents a novel approach that jointly estimates the determinants of sex selective abortions, fertility, and birth spacing, using data on Hindu women from India's National Family and Health Surveys. Women with eight or more years of education in both urban and rural areas are the main users of sex selective abortions and they also have the lowest fertility. Predicted lifetime fertility for these women declined eleven percent between the 1985–1994 and 1995–2006 periods, which correspond to the periods of time before and after sex selection became illegal. Fertility is now around replacement level. This decrease in fertility is accompanied by a six percent increase in the predicted number of abortions during childbearing between the two periods, and sex selection is increasingly used for earlier parities. Hence, the legal steps taken to combat sex selection have been unable to reverse its use. Women with fewer than eight years of education have substantially higher fertility and do not appear to use sex selection.
JEL: J1, O12, I1 Keywords: India, pre-natal sex determination, censoring, competing risk

1 Introduction

During the last century, India has experienced an almost continuous increase in the overall ratio of males to females, mainly because of excess mortality of girls.[1] During the last three decades, the ratio of males to females at birth has also increased as pre-natal sex determination techniques became available, enabling parents to abort fetuses of an unwanted sex.[2] Although our understanding of the determinants of sex selection has improved, there is one important relationship that is still not well understood: the relationship between fertility and sex selection.

Research has been constrained by the absence of direct information on the use of sex selective abortions; few questionnaires ask about sex selection, and those that do show signs of serious under-reporting (Goodkind 1996). In the absence of direct information on sex selective abortions prior research has relied on a simple method for determining the use of sex selection. With data on births one can estimate the effects of various characteristics on the probability of having a son, using the sex of the children born as the dependent variable.[3] In the absence of any interventions the probability of having a son is approximately 0.512 and this probability is independent of genetic factors (Ben-Porath and Welch 1976; Jacobsen, Moller and Mouritsen 1999).[4] With the sex of a fetus random, a statistically significant effect of one or more variables on the probability of giving birth to a boy therefore indicates that sex selection has been used. The main advantage of this method is ease of use; simple OLS or logit can be used to estimate the determinants of the probability that families with a specific set of characteristics will have a boy for a given parity birth.

This method does, however, suffer from two major issues. First, it does not take into account

[1] See Murthi, Guio and Dreze (1995) and Dyson (2001).

[2] See Das Gupta and Bhat (1997), Sudha and Rajan (1999), Arnold, Kishor and Roy (2002), Retherford and Roy (2003) and Jha, Kumar, Vasa, Dhingra, Thiruchelvam and Moineddin (2006). India is not alone; both China and South Korea have seen significant changes in the sex ratio at birth (Zeng, Tu, Gu, Xu, Li and Li 1993; Park and Cho 1995).

[3] Examples of studies that have used this approach are Retherford and Roy (2003), Jha et al. (2006) and Abrevaya (2009).

[4] Poor nutritional and health status of women may lead to relatively more girls being born, probably as an evolutionary response to maximise reproductive success (Trivers and Willard 1973; Wells 2000). Empirical evidence from Ethiopia and from cross-country data supports this hypothesis (Klasen and Wink 2002; Gibson and Mace 2003). See also the discussion for India in Bhaskar and Gupta (2007).

that fertility is a choice—one that is closely related to the use of sex selective abortions. Imagine that families want one son but do not use sex selection. Just below half will not have a son as their first child, by the second child slightly less than 25 percent will not have a son, and so on. If the desire is for one son, more than 99 percent of families would achieve that target if they were willing to have up to six children. If the desire is for one son *and* a maximum of two children, about 25 percent of families have to resort to sex selection to achieve both targets. Hence, in the absence of other changes falling desired fertility increases the use of sex selection. Not only can a decline in desired fertility increase the use of sex selection, increasing the availability of pre-natal sex determination also allows a woman to reach a desired number of sons with fewer births. To fully understand the use of sex selection we therefore need to understand both its use by parity *and* by the fertility decision. This is especially important if lower fertility is the driver of the increased use of sex selection. The simple method does not provide any information on the likelihood of a women progressing to the next pregnancy and therefore cannot tell us how the overall sex ratio changes and what is behind such change.

Secondly, the simple method ignores the fact that spacing between children is an important considerations for many families and that there are significant interactions between birth spacing and sex selection. Each abortion increases the space between births. This additional space can be divided into three parts. First, right after an abortion the uterus needs at least two menstrual cycles to recover before conception should be attempted again because a short (less than three months) space between an abortion and a subsequent pregnancy often leads to a substantial increase in the likelihood of a spontaneous abortion (Zhou, Olsen, Nielsen and Sabroe 2000). Secondly, the expected time to conception is about six months. Finally, reliable sex determination tests can be carried out after three months of gestation. One sex selective abortion therefore delays the next birth by a year on average.[5] If families do not want the space between children to be too long, they may change their decision to use sex selection after one or more abortions and instead give birth to the next child no matter the sex. A similar behavior may also occur if there are concerns about

[5] The waiting time to conception does vary by woman, but even if it is very short, say one month, the minimum additional space between births would be six months per abortion.

possible infertility as a result of many abortions in row without a birth. There is also the possibility that increased access to or knowledge of pre-natal sex determination over time, which can lead women—who, for one reason or another, waited longer to have their next child—to be more likely to use sex selective abortions. The result is that the sample of women who have not yet had a given parity birth may behave differently from the sample of women for whom we can observe the birth. These dynamic selection issues imply that ignoring birth spacing and the potential changes in the use of sex selection with longer spacing will lead the simple method to bias the predicted final sex ratio for a given parity.

This paper presents a novel method that directly incorporates the fact that the use of sex selective abortions affects both the likelihood of having a son *and* the duration between births. The empirical model is a competing risk non-proportional hazard model with two exits states: either a boy or a girl is born. This approach has three major advantages over the simple method. First, it models fertility and spacing decisions jointly with the birth outcome. This allows both progression to next birth and likelihood of using sex selective abortions to be estimated. Secondly, by explicitly incorporating censoring of birth spacing it addresses any potential bias that may arise from changes in the use of sex selection as the duration from the previous birth increases. Finally, even without direct information on the availability or use of pre-natal sex determination, the method can establish what factors determine the use of sex selective abortions because it provides direct information on sex ratios for births at given spell lengths. In sum, for a sample of women the method makes it possible to estimate how many births they will have, what the final sex ratio of their children will be, and how many sex selective abortions they will go through.

Analyzing birth histories from India's National Family and Health Surveys for Hindu women covering the period 1972 to 2006, the method is used to examine three hypotheses that are important for understanding the use of sex selective abortions. First, the reduced demand for children is the main factor in the decision to use sex selective abortions for a given set of son preferences. Secondly, many parents have strong preferences for having only one son, rather than for a large number of sons. Finally, the legal steps taken to combat sex selective abortions have not been able to reverse its use. In addition to understanding the use of sex selective abortions, the results can also provide a solid footing for predicting the long-term impact of the increased use.

The results show that better-educated women in both urban and rural areas are the main users of sex selective abortions and that these women also have the lowest fertility. For women with eight or more years of education and only daughters the use is substantial: Around 60 percent of the children born are boys. As these women's fertility has declined over time, sex selective abortions occur among earlier and earlier parities, although there is no evidence of sex selection on the first birth. Furthermore, the likelihood of having an additional child has declined substantially for women with two children and at least one boy. There is only limited use of sex selective abortions for better-educated women with one or more sons, with the exception of rural women with one son and one daughter, who show evidence of an increased use over time. Women with fewer than eight years of education do not appear to use sex selection and still have relatively high fertility.

There are clear differences between the suggested model and the standard approach for two important cases. The standard approach substantially underestimates the final sex ratio for the third child of urban women with two girls during the initial expansion of access to pre-natal sex determination. On the other hand, the standard method overestimates the final sex ratio for urban women whose first child was a girl during the most recent period, because it ignores the lower use of sex selection for births that are further away from the first birth. Combining the results, completed fertility for well-educated women is predicted to now be around two in urban areas and 2.5 in rural areas. During their childbearing, the number of sex selective abortions one hundred women are predicted to have has risen to 7.8 for urban and 7.3 for rural women. This is a five percent increase in sex selective abortions compared to the period before the practice was made illegal nationwide in India in 1994.

2 Literature Review

The strongest predictor of uneven sex ratio for a given parity is the sex composition of previous children (Retherford and Roy 2003; Jha et al. 2006; Abrevaya 2009). For families without a son, the higher the parity the higher the probability of having a son as the next birth. The propensity to use sex selection increases with socio-economic status, especially education, and the proportion of males to females is larger in cities than in rural areas (Retherford and Roy 2003; Jha et al. 2006). There is substantial disagreement on whether sex selective abortion is used for the first birth. Using the Special Fertility and Mortality Survey Jha et al. (2006) found that in both urban and rural areas there were 54.4 percent boys born among first-borns, while Retherford and Roy (2003) using the first two rounds of the National Family and Health Survey finds little or no evidence of sex selection on the first birth. Although the Special Fertility and Mortality Survey covers a much large number of households than the three rounds of the NFHS combined, Jha et al. (2006) only use births that took place in 1997 making their sample sizes by parity smaller than what is used here. In addition, there are serious questions about the quality of the data, especially the possibility of systematic recall error for girls (George 2006; Bhat 2006).

Despite a large demographic literature on the relation between son preference and fertility stopping behavior (see, for example, Clark 2000), there is little formal analysis of the link between fertility and sex selection (Park and Cho 1995). For India, it has been argued that falling fertility increases the bias against girls (Das Gupta and Bhat 1997), but the stated preference for sons also appear to decline with lower desired fertility (Bhat and Zavier 2003). In Korea simulations suggested that introduction of sex selection changed family size little, but did result in abortions of female fetuses equal to about five percent of actual female births (Park and Cho 1995). For China allowing a three-child policy has been predicted to increase the fertility rate by 35 percent, but also reduce the number of girls aborted by 56 percent (Ebenstein 2011).[6] Most of the sex selection in China is due to parents' with low levels of education (Ebenstein 2011).

[6] This study used data on women who have completed their fertility. No allowance is made for the role of spacing and because only women who have completed their fertility are used, the method cannot be applied to those still in their childbearing years and therefore have censored spells.

Although there is wide agreement that sex selective abortions will substantially affect social and economic development, there is little agreement on the directions of these effects. Girls may ultimately benefit because those born are more likely to be wanted (Goodkind 1996). The reduction in mortality of girls in Taiwan as the number of sex selective abortions have gone up seems to support this hypothesis (Lin, Qian and Liu 2008). A reverse marriage squeeze may also improve conditions for girls (Park and Cho 1995). In India, this would likely result in lower dowries and this may already be happening (Lancaster 2002).

It is, however, not clear that the effects will be uniformly positive. First, the desire for children might fall more rapidly than the total number of desired sons (Das Gupta and Bhat 1997). Secondly, it may result in the development of a female underclass because girls will predominately be born to poorer parents or lead to more crime (Edlund 1999; Edlund, Li, Yi and Zhang 2013). These changes can also affect households' decisions on investment and saving, either because a son has a higher expected return or because parents view savings and investment as a means to improving their son's chances in the marriage market (Ding and Zhang 2009; Wei and Zhang 2009).

3 The Technology of Sex Determination and Selection

There are currently three well-developed technologies for determining the sex of a fetus: Chorionic villus sampling (CVS), amniocentesis and ultrasound. CVS can be applied after the shortest period of gestation (eight to twelve weeks). It is the most complicated but also the most reliable and an abortion can be done in the first trimester. The cost of the procedure appears to vary widely with prices quoted by hospitals on-line from Rs 3,500 to Rs 12,000 (USD 75 to USD 260). Amniocentesis can be performed after fourteen weeks, but requires three to four weeks before the result is available, so an abortion cannot be performed until more than halfway through the second trimester. The cost of amniocentesis is comparable or slightly lower than for CVS. Ultrasound has the advantages of being non-invasive and relatively cheap, from Rs 500 to Rs 1000 (USD 10 to 20). A fetus' sex can be determined in the third month of gestation if it is a boy and the fourth month if it a girl. The first reports of private clinics offering sex determination came in 1982-83 and mobile clinics that can reach remote areas have been available since the mid-1980 in India (Sudha and Rajan 1999).

Abortion has been legal in India since 1971 and still is. Amniocentesis quickly became known as a method for pre-natal sex determination and its use for that purpose became a penal offence. Maharashtra was the first state to pass a law on this in 1988 and in 1994 the Central Government passed the Pre-Natal Diagnostic Techniques (PNDT) Act making determining and communicating the sex of a fetus illegal.[7] There appears to be a substantial leeway in the law, allowing private clinics to operate with little risk of legal action (Sudha and Rajan 1999). The number of convictions under PNDT is very low; in January 2008 Haryana became the first state to reach five convictions.

4 Theory

Consider a T period model where parents decide on fertility sequentially. In each period, parents decide whether to have a pregnancy and whether to pay for pre-natal sex determination given their budget constraint, the characteristics of the children they already have, and the probability of having a son, π.[8] If parents decide to use pre-natal sex determination they automatically abort if the fetus has the unwanted sex at no additional cost. In the next period, parents then decide whether to have another pregnancy and pay for pre-natal sex determination, and so on.

Parents derive utility from the number of boys, b, the number of girls, g, and parental consumption, c, realised at the end of the T periods. The utility function is separable between parental consumption and utility from children. Son preference is defined as higher marginal utility of next child being a boy if parents have an equal number, possibly zero, of boys and girls. A utility function that incorporate son preference is

$$U = u(c) + \alpha u^b(b) + (1 - \alpha) u^g(g). \tag{1}$$

[7] The Act is described in detail at http://pndt.gov.in/.

[8] This model is closely related to the one used by Ejrnæs and Pörtner (2004) to analyse intra-household allocation when fertility is endogenous and parents do not know the characteristics of their next child.

This utility function covers a variety of different preference from equal utility of boys and girls ($\alpha = 0.5$) to an extreme son preference where there is no utility of having a daughter ($\alpha = 1$).

Parents' lifetime income is Y. Parents incur a fixed expense, k, for each child born that covers the basic costs of having a child and the opportunity costs of the mother's time. The cost of girls and boys are assumed to be the same and to ease notation the total number of children in a period is $n = b + g$. Pre-natal sex determination carries a cost p_s, and the number of pregnancies where sex determination has taken place is s. The budget constraint is

$$c + nk + sp_s \leq Y. \tag{2}$$

The decision tree for a model with just two periods is shown in Figure 1 and illustrate the complexity of the optimisation problem. In the first period, parents can either have no children, have a child without using pre-natal sex determination ($-S$) or use pre-natal sex determination (S). In the latter case, parents enter the second period with either a boy or no children. This decision then repeats itself in the second period (except for those who decided not to have a child in the first period). Looking only at those with two pregnancies there are now nine possible outcomes rather than three if pre-natal sex determinations were not available.

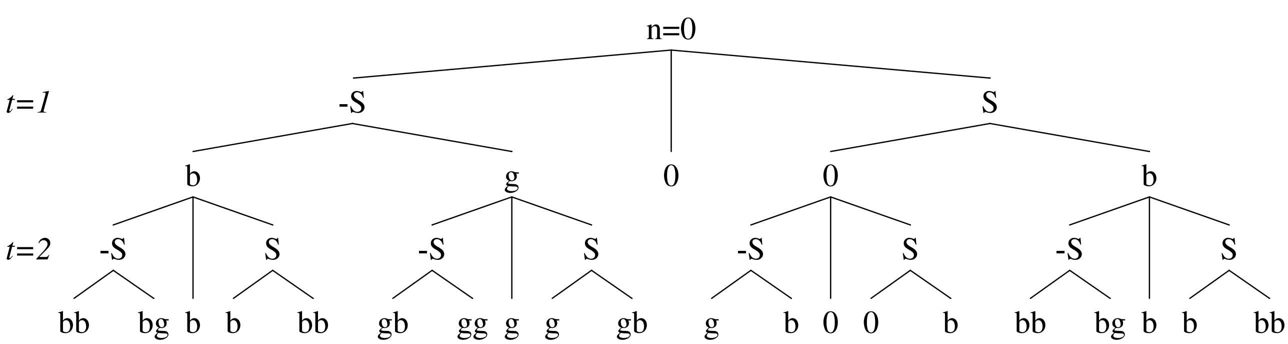

Figure 1: Decision tree for a two period model with pre-natal sex screening

In principle this problem can be solved using dynamic programming, but because the choice variable and the state variables are discrete a closed-form solution is generally not available. There are three discrete state variables (boys, girls, scans) and one choice variable with three choices

Table 1: Fertility and Percent Boys by Cost of Children and Pre-Natal Sex Determination

Cost of children (k)	Cost of Pre-Natal Sex Determination (p_s) 1			3			5			No Access	
	Fertility			Fertility			Fertility			Fertility	
	Mean	St. Dev.	Percent Boys[a]	Mean	St. Dev.	Percent Boys[a]	Mean	St. Dev.	Percent Boys[a]	Mean	St. Dev.
				Son Preference $\alpha = 0.55$							
12	4.09	0.38	59.03	4.40	0.55	50.71	4.45	0.66	50.00	4.45	0.66
14	3.68	0.58	60.13	3.68	0.58	51.74	3.75	0.66	50.07	3.75	0.66
16	3.12	0.33	54.04	3.12	0.33	54.04	3.25	0.43	50.07	3.25	0.43
18	2.25	0.43	61.04	2.50	0.50	54.99	2.50	0.50	49.99	2.50	0.50
				Son Preference $\alpha = 0.65$							
12	4.49	0.51	63.06	4.34	0.48	54.68	4.42	0.61	50.15	4.43	0.65
14	3.12	0.34	69.94	3.47	0.50	58.58	3.68	0.58	50.91	3.71	0.67
16	3.00	0.06	70.79	3.00	0.02	54.22	3.12	0.33	52.03	3.19	0.52
18	2.75	0.43	72.64	2.00	0.02	62.39	2.25	0.43	55.48	2.37	0.70
				Son Preference $\alpha = 0.75$							
12	3.98	0.17	76.41	4.00	0.06	59.30	4.50	0.79	50.39	4.52	0.87
14	3.84	0.37	77.21	3.00	0.05	70.78	3.50	0.50	51.86	3.61	0.76
16	3.00	0.06	70.79	2.97	0.18	70.49	3.00	0.02	54.22	3.19	0.52
18	2.00	0.05	99.99	2.32	0.47	67.55	2.00	0.02	62.39	2.37	0.70

Note. Simulations based on 50,000 "individuals", 12 periods, $Y = 200$, probability of a son of 0.5.
[a] Percent boys is the percent boys of all births.

(no pregnancy, pregnancy without pre-natal sex determination, and pregnancy with pre-natal sex determination). Combined with the finite time horizon a problem of this type is especially suited for simulation and the model is simulated using backward induction.

The utility function is

$$U = 1.5\ln(c-30) + \alpha\ln(b+1) + (1-\alpha)\ln(g+1).$$

All simulations assume 12 periods, $Y = 200$ and the probability of a son equal to 0.5. The three "exogenous" variables and their examined values are the cost of children ($k \in 12,14,16,18$), cost of pre-natal sex determination ($p \in 1,3,5,30$), where the last value is equivalent to no pre-natal sex determination available, and the degree of son preference ($\alpha \in 0.55,0.65,0.75$). The optimal choice sets are applied to a sample of 50,000 "individuals", who each have 12 "potential" children, whose sex come from a binomial distribution with the probability above.

Table 1 shows average fertility, the standard deviation of fertility and the ratio of boys to births for selected values of son preference (α), cost of pre-natal sex determination (p_s) and cost of children (k). In the absence of pre-natal sex determination, stronger son preference is often thought to increase fertility. This is based on a confusion of total fertility with parity progression decisions based on sex composition. For a given parity level, parents with a relatively large number of girls are more likely to have another child, but this effect is countered by a lower likelihood of progression for those with relatively more sons. Except for when the cost of children is low, average fertility falls with higher son preference. An often ignored aspect of the impact of son preference is the effect on the variance of fertility. Stronger son preference should increase the variance in completed fertility; families who have sons early also stop childbearing early, while those with girls continues for much longer. The simulations show that the variance in fertility increases with stronger son preference, except when children are very expensive.

Introducing pre-natal sex determination has two opposing effects on fertility. On one hand, parents can reach a desired number of sons with fewer births and this would tend to reduce fertility. On the other hand, parents will not have to pay for unwanted children and this wealth effect will tend to increase fertility. The simulations show that reducing the cost of pre-natal sex determination tend to reduce fertility (and lower its variance), but the small reductions are consistent with the two opposing effects being of approximately equal size.[9] As sex selection becomes less expensive the proportion of boys clearly increases and, correspondingly, the spells between births become longer. In addition, the higher the parity the longer the duration between births.

An important point is that a household may reverse its decision to use sex selection within a spell. Consider a household in the last period that has been using sex selection in previous periods: if this household derives higher utility from having a girl than paying for pre-natal sex determination and possibly not having a child, the optimal decision is not to use sex selection, even if abortions were used on previous pregnancies. Other factors, such as dis-utility of too long spacing between births, not captured by the model can have the same effect. Dis-utility of long spac-

[9] See also the discussions in Leung (1994) and Davies and Zhang (1997).

ing might arise simply from wanting children close in age, but could also come from decreasing economies of scale when children are far from each other in age or from larger opportunity costs of leaving the labour market at older ages. Concerns about possible infertility may also increase with the number of abortions. Hence, households may trade off between their preference for a son and their preference for shorter birth spacing. The change in use of pre-natal sex determination after abortions is clearly more likely to happen the fewer children a family has; if the household is, for example, willing to have three children there may to room to change the decision on sex selection within a spell for the second child.

The more households change their use of pre-natal sex determination within a spell, the more likely it is that the simple method produces biased results. On one hand, if we observe all births a set of women will have, the standard method may provide a downward biased estimate of the use of sex selection. On the other hand, if many birth spells are censored the predicted ratio of boys to girls will be biased upward because those who have not had a birth yet are more likely to have a girl. In addition, with censoring it will appear as if fertility is lower than it really is, although the standard method obviously cannot estimate whether a women will have a birth.

The model informs two of the hypotheses to be examined. The first is the "fertility" hypothesis: The lower demand for children is the main factor in the decision to use sex selective abortions for a given set of son preferences. Two important factors in predicting fertility are education and area of residence; the more educated a women is and the more urbanised the areas she lives in, the lower is her fertility. Hence, the hypothesis can be examined by comparing the use of sex selection between education levels and between urban and rural areas.

The second hypothesis tests whether the differential stopping behavior observed in many studies is an indication that families have preferences for one son (Repetto 1972; Arnold, Choe and Roy 1998; Dreze and Murthi 2001). The Hindu tradition requires a son for lighting the funeral pyre and, hence, this can be called the "funeral pyre" hypothesis (Arnold et al. 1998). If a family has at least one boy there should be no evidence of sex selective abortions if the "funeral pyre" hypothesis holds. If sex selection is instead decreasing in the relative number of sons it is more likely that parents have a set of preferences as in the model above.

5 Estimation Strategy

The theoretical model has three implication for parents' fertility decisions when there is access to sex selective abortions: a higher probability that the next child is a son, longer average waiting time to next birth, and that the use of pre-natal sex determination is not necessarily constant between pregnancies within a spell (between one observed birth and the next). All three must be incorporated into the estimation strategy to provide unbiased and precise estimates of the use of sex selective abortions. This section discusses the empirical specifications.

The empirical model is a discrete time non-proportional competing risk hazard model with two exits states: either a boy or a girl is born.[10] For each married woman, $i = 1, \ldots, n$, in the data we observe at least one spell, with all spell durations measured in quarters. The first spell begins at the time of marriage and subsequent spells nine months after the birth of the previous child. The first spell begins at marriage because a number of women report giving birth less than nine months after they are married. The second and subsequent spells begin at nine months after the previous birth because that is the earliest we would expect to observe a new birth. In other words, for women who have had one or more births, their second spell begins nine months after the birth of the first child and continues until either the second child is born or the survey takes place. The starting point for each spell is $t = 1$ and it continues until time t_i when either a birth occurs or the survey takes place (the observation is censored). The time of censoring is assumed independent of the hazard rate as is standard in the literature.

There are two exit states: birth of a boy, $j = 1$, or birth of a girl, $j = 2$, and J_i is a random variable indicating which event took place. The discrete time hazard rate h_{ijt} for an event is[11]

$$h_{ijt} = \Pr(T_i = t, J_i = j \mid T_i \geq t; \mathbf{Z}_{it}, \mathbf{X}_i), \tag{3}$$

[10] Merli and Raftery (2000) used a discrete hazard model to examine whether there were under-reporting of births in rural China, although they estimated separate waiting time regressions for boys and girls.

[11] To ease presentation the indicator for spell number is suppressed.

where T_i is a discrete random variable that captures the quarter in which a birth for woman i occurs. The vectors of explanatory variable $\mathbf{Z}_{it}$ and $\mathbf{X}_i$ includes information about various individual, household, and community characteristics discussed below.

The hazard rate is specified as

$$h_{ijt} = \frac{\exp(D_j(t) + \alpha'_{jt}\mathbf{Z}_{it} + \beta'_j\mathbf{X}_i)}{1 + \sum_{l=1}^{2} \exp(D_j(t) + \alpha'_{lt}\mathbf{Z}_{it} + \beta'_l\mathbf{X}_i)} \quad j = 1,2 \tag{4}$$

where $D_j(t)$ is the piece-wise linear baseline hazard for outcome j, captured by dummies and the associated coefficients,

$$D_j(t) = \gamma_{j1}D_1 + \gamma_{j2}D_2 + \ldots + \gamma_{jT}D_T, \tag{5}$$

where $D_m = 1$ if $t = m$ and zero otherwise.[12] This approach to modeling the baseline hazard is flexible and does not place overly strong restrictions on the baseline hazard.

Specifying the model as a proportional hazard model, i.e. one where explanatory variables simply shift the hazard rates up or down independent of spell length, would in principle result in greater efficiency, provided that the proportionality assumption holds. The problem here is that the proportional hazard model does not allow for differences in effects of covariates over time within a spell and ignoring differences in the shape of the hazard functions between different types of individuals can lead to substantial bias. It is highly unlikely, even in the absence of pre-natal sex determination, that the baseline hazards are the same across education levels, areas of residence or sex composition of previous births. The bias from the proportionality assumption is likely to be exacerbated by the introduction of pre-natal sex determination. This is especially the case if the use of sex selection within a spell responds to the length of the spell and that is one of the major points of the theoretical model.

The model, therefore, is specified as a non-proportional model where the main explanatory variables and the interactions between them are interacted with the baseline hazards. This is cap-

[12] The exact period covered by the individual Ds depends on the amount of available information for a spell but are quarters whenever possible.

tured by the Z set of explanatory variables

$$\mathbf{Z}_{it} = D_j(t) \times (\mathbf{Z}_1 + Z_2 + \mathbf{Z}_1 \times Z_2), \tag{6}$$

where $D_j(t)$ is the piece-wise linear baseline hazard and $\mathbf{Z}_1$ captures sex composition of previous children, if any, and Z_2 area of residence. This allows the effects on the probabilities of having a boy, a girl, or no birth to vary over time within a spell for the main variables. To further minimise the potential bias, estimations are done separately for different levels of the mother's education and for different time periods as described below. The remaining explanatory variables, **X**, enter proportionally.[13] The use of a non-proportional specification together with a flexible baseline hazard is also important because potential effects of unobserved heterogeneity are mitigated the more flexible the baseline hazard specification (Dolton and von der Klaauw 1995).

Equation (4) is equivalent to the logistic hazard model and has the same likelihood function as the multinomial logit model (Allison 1982; Jenkins 1995). Hence, if the data are transformed so the unit of analysis is spell unit rather than the individual woman, the model can be estimated using a standard multinomial logit model.[14] In the reorganized data the outcome variable is zero if the woman does not have a child in a given period, one if she gives birth to a son in that period, and two if she gives birth to a girl in that period.

Direct interpretation of the estimated coefficients for this model is challenging because of the competing risk setup. First, coefficients show the change in hazards relative to the base outcome, no birth, rather than simply the hazard of an event. Second, a positive coefficient does not necessarily imply that an increase in the value of a variable increases the probability that the associated event occurs because the probability of another event(s) may increase even more (Thomas 1996). The

[13] It is possible to additionally incorporate time-variant characteristics of the explanatory variables in this set-up, but here none of the explanatory variable change within a spell in the data.

[14] A potentially issue is that the multinomial model assumes that alternative exit states are stochastically independent, also known as the Independence of Irrelevant Alternatives (IIA) assumption. This assumption rules out any individual-specific unmeasured or unobservable factors that affect both the hazard of having a girl and the hazard of having a boy. To address this issue the estimations include a proxy for fecundity discussed in Section 6. In addition, the multivariate probit model can be used as an alternative to the multinomial logit because the IIA is not imposed (Han and Hausman 1990). The results are essentially identical between these two models and available upon request.

model does, however, make it straightforward to calculate the predicted probabilities of having a boy and of having a girl for each quarter, conditional on a set of explanatory variables and not having had a child before that quarter. The predicted probability of having a boy in quarter t for a given set of explanatory variable values, $\mathbf{Z}_k$ and $\mathbf{X}_k$, is

$$P(b_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t) = \frac{\exp(D_j(t) + \alpha_{1t}'\mathbf{Z}_{kt} + \beta_1'\mathbf{X}_k)}{1 + \sum_{l=1}^{2} \exp(D_j(t) + \alpha_{lt}'\mathbf{Z}_{kt} + \beta_l'\mathbf{X}_k)}, \tag{7}$$

and the predicted probability of having a girl is

$$P(g_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t) = \frac{\exp(D_j(t) + \alpha_{2t}'\mathbf{Z}_{kt} + \beta_2'\mathbf{X}_k)}{1 + \sum_{l=2}^{2} \exp(D_j(t) + \alpha_{lt}'\mathbf{Z}_{kt} + \beta_l'\mathbf{X}_k)}. \tag{8}$$

With these two probabilities is it easy to calculate, for each quarter, the predicted percentage of children born that are boys and the associated confidence interval for given values of explanatory variables,

$$\hat{Y}_t = \frac{P(b_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t)}{P(b_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t) + P(g_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t)} \times 100. \tag{9}$$

For ease of exposition the procedure is presented here in two steps but the actual calculation of the percent boys is done in one step with the 95 percent confidence interval calculated using the Delta method. Results are presented as the predicted percent boys born by length of birth spacing using graphs.[15] For each graph, the extent to which the predicted percent boys is statistically significantly above the natural sex ratio then indicate the extent of the use of sex selection.

Another important aspect of the hazard model is the survival curve, which shows the predicted probability of not having had a birth yet by spell duration. The survival curves are presented for two reasons. First, they show the predicted progression to the next birth by quarter and how quickly it occur, if at all. Secondly, they provide "weighting" for the associated predicted percentage boys born. The steeper the survival curve, the more weight should be assigned to a given spell period because it is based on more births, whereas if the survival curve is flat the period should be given

[15] The parameter estimates are available on request.

little weight because the predicted percentage boys is based on few births. The survival curve for quarter t is

$$S_t = \prod_{d=1}^{t} \left(1 - \left(P(b_d|\mathbf{X}_k, \mathbf{Z}_{kd}, d) + P(g_d|\mathbf{X}_k, \mathbf{Z}_{kd}, d)\right)\right), \tag{10}$$

or equivalently

$$S_t = \prod_{d=1}^{t} \left(\frac{1}{1 + \sum_{l=2}^{2} \exp(D_j(t) + \alpha'_{ld}\mathbf{Z}_{kd} + \beta'_l \mathbf{X}_k)} \right). \tag{11}$$

6 Data

The data come from the three rounds of the National Family Health Survey (NFHS-1, NFHS-2 and NFHS-3), collected in 1992-93, 1998-99 and 2005-2006, respectively.[16] They were collected by the International Institute for Population Sciences in Mumbai and have nationwide coverage. There are three advantages to using the NFHS. First, the data are considered high quality. There still is a potential issue of recall error, but, as shown below, it is more easily addressed than in the data used by Jha et al. (2006). Secondly, it has complete birth histories for a large number of women. Finally, by combining the three NFHS rounds it is possible to show the development from before sex selective abortions were available until 2006. Specifically, it is possible to validate the method using data before sex selection and to show whether the law banning providing information about the sex of a fetus have changed the use of sex selective abortions.

The surveys are large: NFHS-1 covered 89,777 ever-married women aged 13-49 from 88,562 households, NFHS-2 covered 90,303 ever-married women aged 15-49 from 92,486 households and NFHS-3 covered 124,385 never-married and ever-married women aged 15-49 from 109,041 households. Visitors to the household were dropped from the sample as were women married more than once, and women with inconsistent information on age of marriage or missing information on education. Women interviewed in NFHS-3 who were never married or were gauna had not been performed were also dropped. Women who had at least one multiple birth, reported having a birth before age 12, had a birth before marriage or a duration between births less than nine months were

[16] NFHS-2 also has a small number of observation collected in 2000, due to a delay in the survey for Tripura.

dropped. There are a number of women for whom the space between marriage and first birth were less than nine months. These women are in the sample unless they are dropped for any of the other reasons. Finally, only Hindu women are used here. About 80 percent of India's population are Hindu and focusing on this group ensures fewer issues with estimating the hazards functions. Specifically, it avoids imposing the assumption that the baseline hazard is the same for Hindu women and other women, such as Sikhs, which would be problematic if use of sex selection varies across groups.[17]

6.1 Recall Error

A data issue that has not received the attention it deserves is the reliability of the birth histories. Of primary concern is the potential for systematic recall error, i.e. children who are missed non-randomly, for example, because of those children's early death. Interviewers therefore probe for any missed births if there is four or more years between two births reported as consecutive. Although probing catches many missed births, recall error is still likely to be a substantial problem in India. First, there is a significantly higher mortality risk for girls than for boys. Secondly, the preference for boys also lead to more boys than girls being remembered among those children who died. Finally, the probing works of long intervals between observed births, but given short durations between births, especially after the birth of a girl, it is unlikely to pick up all children.

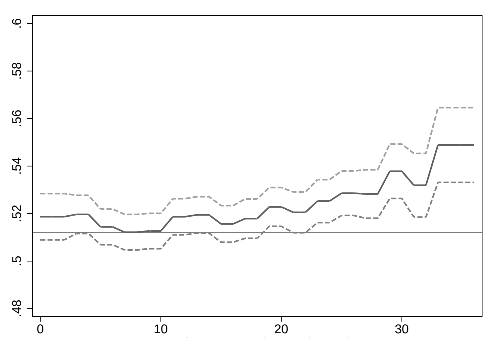

Figure 2: Ratio of boys in "first" births

[17] Furthermore, the sample sizes for other groups, such as Sikhs, are small relative to Hindu, potentially introducing excessive noise in the results.

Figure 2 shows that differential recall of boys and girls is, indeed, a problem. The solid line is the ratio of boys to children reported as first born by the number of years between the survey and marriage, the dashed lines indicate the 95 percent confidence interval and the horizontal line the natural ratio of boys (approximately 0.512).[18] The observed ratio of boys is increasingly above the expected value the longer ago the parents were married. Pre-natal sex determination techniques did not become widely available until the mid-eighties and therefore cannot explain the higher sex ratio. A more likely explanation is recall error, with the first child being a girl who died early, followed by a son who is then reported as being the first born. The increasingly unequal sex ratio with increasing marriage duration suggests that a simple solution to the recall error problem is to drop observations for women who were married "too far" from the survey year. The cut-off points used vary by survey round. Women who were married 22 years or more before the their interview date in NFHS-1 were dropped and the corresponding cut-off points for NFHS-2 and NFHS-3 are 23 and 26 years.[19] The final sample consists of 115,210 Hindu women.

6.2 Variables

The dependent variables are duration of the spells and outcomes of those spells. A spell has three possible outcomes: the birth of a boy, the birth of a girl or the spell is censored. The first spell is the duration from marriage until the first birth (or until censoring occur). Because a number of women married early (and began living with their husbands at a young age) the exact starting point of this spell present some difficulties. For those who marry young, the "correct" starting point is first ovulation, when the woman becomes "at risk" for a birth, but often menarche is used instead. Information on age of menarche is only available in NFHS-1. Instead, for women that began living with their husband before age 12, the staring point is the month they turned 12 years old.

The second spell is from nine months after the birth of the first child until the second child

[18] To ensure sufficient cell sizes the years are grouped in twos.

[19] The graph for second births shows a similar pattern with the likelihood of the second child being a boy going up with increasing marriage duration. The graphs for the second births and the individual survey rounds are available upon request.

is born or censoring occur. The third and fourth spells follow the same definition. Beginning the spell nine months after the previous birth allows the next pregnancy to be carried to term. In other words, there is no variation in the dependent variable for the nine months following a prior birth.[20]

Because the method requires births to occur, spells are counted as censored once data become too sparse for the method to work. For the first and second spell an observation is counted as censored if a woman has not had child within 21 quarters of the beginning of the spell and whereas the cut-off is 19 quarters for the third and fourth spell. This is equivalent to 6 years from last birth or marriage for the first and second spells and 5.5 years from last birth for third and fourth spell. The shorter spell length for the third and fourth spell is dictated by the low number of births after 19 quarters. Sterilization is also treated as censored.

The explanatory variables are divided into two groups. The first group consists of the main explanatory variables that are expected to affect the shape of the hazard function and includes the mother's education, the sex composition of previous children and the area of residence. The second group consists of the variables that are expected to have a proportional effect on the hazard and includes the age of the mother at the beginning of the spell, the length of her first spell (for second spell and above), whether the household owns land, and whether the household belongs to a scheduled tribe or caste. There is clearly a trade-off between increasing the number of variables that are expected to affect the shape of the hazard and the precision of the results. The more variables are interacted with the baseline hazard the more data is required to precisely estimate the effects. The choice of the variables to interact with the baseline hazard is based on which variables are likely to be most important for fertility decisions and the use of sex selection.

The mother's education is a major determinant of the opportunity cost of time and higher education is expected to lead to lower fertility, thereby increasing the use of sex selective abortions. Father's education had little effect on the hazards and the use of sex selective abortions and is not included. Education is divided into three groups: No education, one to seven years of education, and eight and more years of education. The models are estimated separately for each education

[20] There are a few women that report births that occur less than nine months after the previous birth. Those women are dropped from the sample.

level.

The sex composition of previous children affects both the timing of fertility and the use of sex selective abortions. The exact effects depend on whether the "funeral pyre" hypothesis holds and whether parents continue to try to conceive sooner after a girl than a boy after pre-natal sex determination become available. Sex composition is captured by dummy variables for the possible combinations for the specific spell, ignoring the ordering of births. As an example, for the third spell three groups are used: Two boys, one girl and one boy, and two girls.

The area of residence is a dummy variable for the household living in an urban area.[21] Area of residence captures both access to pre-natal sex determination techniques and the cost of children. Moving from rural to more urban area should increase both the cost of children and access to pre-natal sex determination. Both are expected to lead to greater use of sex selective abortions.

The sex composition of children, area of residence and their interactions are interacted with the piece-wise linear baseline hazard dummies. In other words, the baseline hazards are assumed to be different depending on where a woman lives and the composition of her previous children. As an example, for the second spell a separate regression is run for each education level and in each regression four different baseline hazards are estimated (first child a boy in rural area, first child a boy in urban area, first child a girl in rural area, first child a girl in urban area). Although this substantially increases the number of regressions and estimated parameters it reduces the potential problem of including other variables as proportional effects. In addition, it forces attention to how much data is really available to estimate the determinants of sex selective abortions.

The remaining variables are expected to affect hazards proportionally. Although fecundity cannot be observed directly a suitable proxy is the duration from marriage until first birth. Most Indian women do not use contraception before the first birth and there is pressure to show that a newly married woman can conceive. This is confirmed by the very short spells between marriage and first birth, even among the most educated. Hence, long spell between marriage and first birth is likely due to low fecundity. The expected effect of a longer first spell is to reduce sex selective

[21] NFHS uses four categories for area of residence: Large city, small city, town and countryside. To reach a sufficient sample size urban areas are merged into one group.

abortions in subsequent spells. For both this variable and the age of the mother at the beginning of the spell the squares are also included. The remaining variables are dummies for household ownership of land and membership of a scheduled caste or tribe.

6.3 Descriptive Statistics

Tables 2 and 3 present descriptive statistics for the four spells by education level and when the spell began. Three time periods are covered: 1972-1984, 1985–1994 and 1995–2006. A spell belongs to a given period if the birth or marriage beginning the spell took place in that period. The first period covers before sex selective abortions became widely available. Although abortion was legalised in 1971 and amniocentesis was introduced in India in 1975, the first newspaper reports of availability of sex selection techniques are in 1982-83 (Sudha and Rajan 1999). The number of clinics quickly increased and knowledge became widespread after a senior official's wife aborted a fetus that turned out to be male (Sudha and Rajan 1999, p. 598). The second period covers the time from the widespread emergence of sex selective abortion until the Pre-natal Diagnostic Techniques (PNDT) act was passed in 1994. The final period is from the PNDT act until the last available survey. The PNDT made is a criminal offence to reveal the sex of the fetus and was followed by a campaign against the use of sex selection, although enforcement appears to be relatively lax. Dividing spells into the three period allows for an analysis of how fertility decisions and the use of sex selection has changed over time under three different regimes. For each spell the three first variable are the possible outcomes of the spell: Boy, girl or censored. These are followed by the sex composition of previous children (for spells two to four) and other explanatory variables. Finally, the last two rows show the number of quarters observed and the number of women who began the spell. Within each spell, each row is a separate regression sample used for the results below.

There is a substantial number of censored observations. As an example, for highly educated women who had their first child in the 1995–2006 period, almost half did not have their second child by the time of survey. Hence, although about thirteen thousand women began the second spell there is only about seven thousand births to these women. Censoring becomes even more important

Table 2: Descriptive Statistics by Education Level and Beginning of Spell for First and Second Spell

		No Education			1-7 Years of Education			8+ Years of Education		
		1972-1984	1985–1994	1995–2006	1972-1984	1985–1994	1995–2006	1972-1984	1985–1994	1995–2006
First Spell	Boy born	0.458	0.422	0.363	0.461	0.436	0.388	0.472	0.447	0.404
		(0.498)	(0.494)	(0.481)	(0.498)	(0.496)	(0.487)	(0.499)	(0.497)	(0.491)
	Girl born	0.418	0.401	0.355	0.430	0.419	0.368	0.449	0.421	0.383
		(0.493)	(0.490)	(0.478)	(0.495)	(0.493)	(0.482)	(0.497)	(0.494)	(0.486)
	Censored	0.124	0.176	0.282	0.109	0.144	0.244	0.079	0.132	0.213
		(0.329)	(0.381)	(0.450)	(0.311)	(0.351)	(0.429)	(0.270)	(0.339)	(0.409)
	Urban	0.175	0.170	0.206	0.359	0.333	0.335	0.693	0.612	0.573
		(0.380)	(0.376)	(0.404)	(0.480)	(0.471)	(0.472)	(0.461)	(0.487)	(0.495)
	Age	15.813	16.225	16.780	16.888	17.387	17.732	19.453	19.990	20.716
		(2.423)	(2.633)	(2.896)	(2.719)	(3.054)	(3.320)	(3.360)	(3.663)	(3.950)
	Owns land	0.592	0.587	0.543	0.507	0.500	0.471	0.319	0.384	0.404
		(0.491)	(0.492)	(0.498)	(0.500)	(0.500)	(0.499)	(0.466)	(0.486)	(0.491)
	Scheduled caste	0.214	0.236	0.271	0.126	0.169	0.234	0.052	0.092	0.135
		(0.410)	(0.425)	(0.445)	(0.332)	(0.375)	(0.423)	(0.223)	(0.288)	(0.342)
	Scheduled tribe	0.137	0.154	0.164	0.043	0.054	0.080	0.016	0.026	0.038
		(0.344)	(0.361)	(0.370)	(0.204)	(0.226)	(0.271)	(0.126)	(0.158)	(0.190)
	Number of quarters	224,978	170,869	47,701	72,008	68,772	32,879	63,719	93,262	65,376
	Number of women	24,460	21,536	7,330	8,779	10,053	5,798	9,038	15,119	13,097
Second Spell	Boy born	0.489	0.444	0.361	0.475	0.435	0.340	0.456	0.399	0.288
		(0.500)	(0.497)	(0.480)	(0.499)	(0.496)	(0.474)	(0.498)	(0.490)	(0.453)
	Girl born	0.450	0.412	0.327	0.461	0.412	0.320	0.427	0.352	0.242
		(0.498)	(0.492)	(0.469)	(0.498)	(0.492)	(0.467)	(0.495)	(0.478)	(0.429)
	Censored	0.061	0.144	0.313	0.064	0.153	0.340	0.117	0.248	0.470
		(0.240)	(0.351)	(0.464)	(0.245)	(0.360)	(0.474)	(0.322)	(0.432)	(0.499)
	One boy	0.527	0.515	0.507	0.519	0.514	0.504	0.513	0.520	0.515
		(0.499)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
	One girl	0.473	0.485	0.493	0.481	0.486	0.496	0.487	0.480	0.485
		(0.499)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
	Urban	0.183	0.181	0.213	0.370	0.350	0.347	0.695	0.628	0.580
		(0.386)	(0.385)	(0.409)	(0.483)	(0.477)	(0.476)	(0.460)	(0.483)	(0.494)
	Age	17.786	18.356	18.872	18.637	19.208	19.576	21.007	21.597	22.258
		(2.746)	(3.063)	(3.301)	(2.873)	(3.228)	(3.382)	(3.386)	(3.630)	(3.910)
	First spell length	24.644	28.618	29.267	20.973	24.117	24.835	19.416	21.554	22.259
		(20.051)	(24.986)	(27.439)	(17.803)	(22.119)	(23.391)	(15.039)	(18.570)	(19.265)
	Owns land	0.587	0.569	0.525	0.493	0.488	0.465	0.313	0.368	0.400
		(0.492)	(0.495)	(0.499)	(0.500)	(0.500)	(0.499)	(0.464)	(0.482)	(0.490)
	Scheduled caste	0.215	0.241	0.269	0.124	0.169	0.231	0.053	0.085	0.134
		(0.411)	(0.428)	(0.443)	(0.330)	(0.375)	(0.422)	(0.223)	(0.279)	(0.341)
	Scheduled tribe	0.137	0.149	0.161	0.042	0.051	0.073	0.015	0.026	0.035
		(0.344)	(0.356)	(0.368)	(0.201)	(0.220)	(0.260)	(0.121)	(0.159)	(0.183)
	Number of quarters	170,541	227,151	66,693	65,653	102,392	48,186	75,122	161,610	117,022
	Number of women	21,052	28,393	9,442	8,187	12,566	6,556	8,156	16,612	13,527

Note. Means without parentheses and standard deviation in parentheses. Interactions between variables, baseline hazard dummies and squares not shown.

Table 3: Descriptive Statistics by Education Level and Beginning of Spell for Third and Fourth Spell

		No Education			1-7 Years of Education			8+ Years of Education		
		1972-1984	1985–1994	1995–2006	1972-1984	1985–1994	1995–2006	1972-1984	1985–1994	1995–2006
Third Spell	Boy born	0.478	0.426	0.322	0.454	0.388	0.263	0.352	0.267	0.165
		(0.500)	(0.495)	(0.467)	(0.498)	(0.487)	(0.440)	(0.478)	(0.442)	(0.372)
	Girl born	0.441	0.395	0.309	0.421	0.358	0.242	0.314	0.216	0.130
		(0.497)	(0.489)	(0.462)	(0.494)	(0.480)	(0.428)	(0.464)	(0.412)	(0.336)
	Censored	0.081	0.179	0.369	0.125	0.253	0.495	0.334	0.517	0.705
		(0.273)	(0.384)	(0.483)	(0.331)	(0.435)	(0.500)	(0.472)	(0.500)	(0.456)
	Two boys	0.276	0.257	0.241	0.247	0.245	0.230	0.265	0.242	0.236
		(0.447)	(0.437)	(0.428)	(0.431)	(0.430)	(0.421)	(0.441)	(0.428)	(0.425)
	One boy, one girl	0.490	0.502	0.506	0.511	0.500	0.516	0.489	0.511	0.527
		(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.499)
	Two girls	0.235	0.241	0.253	0.242	0.255	0.254	0.246	0.247	0.237
		(0.424)	(0.428)	(0.435)	(0.429)	(0.436)	(0.435)	(0.431)	(0.432)	(0.425)
	Urban	0.181	0.183	0.209	0.378	0.357	0.344	0.691	0.646	0.570
		(0.385)	(0.387)	(0.407)	(0.485)	(0.479)	(0.475)	(0.462)	(0.478)	(0.495)
	Age	20.008	20.696	21.273	20.841	21.453	21.986	22.988	23.832	24.551
		(2.892)	(3.191)	(3.545)	(2.943)	(3.240)	(3.620)	(3.417)	(3.751)	(4.075)
	First spell length	23.419	27.047	27.243	19.866	23.406	23.558	18.662	20.570	21.245
		(18.476)	(22.661)	(24.513)	(16.433)	(20.188)	(21.225)	(14.289)	(16.660)	(17.537)
	Owns land	0.599	0.573	0.537	0.501	0.501	0.488	0.322	0.368	0.418
		(0.490)	(0.495)	(0.499)	(0.500)	(0.500)	(0.500)	(0.467)	(0.482)	(0.493)
	Scheduled caste	0.208	0.244	0.272	0.118	0.168	0.231	0.053	0.083	0.136
		(0.406)	(0.429)	(0.445)	(0.322)	(0.374)	(0.422)	(0.225)	(0.275)	(0.342)
	Scheduled tribe	0.136	0.148	0.160	0.043	0.052	0.068	0.013	0.026	0.034
		(0.342)	(0.355)	(0.367)	(0.202)	(0.222)	(0.252)	(0.115)	(0.158)	(0.182)
	Number of quarters	104,473	201,194	68,155	41,929	85,630	40,334	48,441	124,339	77,410
	Number of women	12,976	25,020	9,295	4,890	9,752	4,997	4,257	10,602	7,836
Fourth Spell	Boy born	0.464	0.387	0.295	0.402	0.350	0.235	0.310	0.229	0.173
		(0.499)	(0.487)	(0.456)	(0.490)	(0.477)	(0.424)	(0.463)	(0.420)	(0.378)
	Girl born	0.409	0.366	0.271	0.392	0.305	0.200	0.297	0.197	0.115
		(0.492)	(0.482)	(0.444)	(0.488)	(0.460)	(0.400)	(0.457)	(0.398)	(0.319)
	Censored	0.127	0.247	0.434	0.206	0.345	0.566	0.393	0.574	0.713
		(0.333)	(0.431)	(0.496)	(0.405)	(0.475)	(0.496)	(0.489)	(0.495)	(0.453)
	Three boys	0.136	0.123	0.111	0.107	0.106	0.101	0.107	0.102	0.082
		(0.343)	(0.328)	(0.315)	(0.309)	(0.307)	(0.301)	(0.309)	(0.302)	(0.274)
	Two boys, one girl	0.370	0.358	0.345	0.345	0.320	0.307	0.354	0.303	0.288
		(0.483)	(0.479)	(0.475)	(0.476)	(0.466)	(0.461)	(0.478)	(0.460)	(0.453)
	One boys, two girls	0.365	0.390	0.400	0.405	0.414	0.424	0.395	0.440	0.468
		(0.481)	(0.488)	(0.490)	(0.491)	(0.493)	(0.494)	(0.489)	(0.496)	(0.499)
	Three girls	0.129	0.130	0.143	0.143	0.160	0.168	0.145	0.156	0.162
		(0.335)	(0.336)	(0.350)	(0.350)	(0.367)	(0.374)	(0.352)	(0.363)	(0.369)
	Urban	0.172	0.182	0.198	0.361	0.347	0.325	0.667	0.599	0.513
		(0.378)	(0.385)	(0.398)	(0.480)	(0.476)	(0.468)	(0.471)	(0.490)	(0.500)
	Age	21.958	22.887	23.607	22.642	23.520	24.307	24.253	25.480	26.067
		(3.011)	(3.332)	(3.708)	(2.889)	(3.359)	(3.754)	(3.193)	(3.706)	(4.141)
	First spell length	22.063	26.033	26.624	19.239	22.523	24.255	17.701	20.131	21.480
		(17.051)	(21.346)	(23.462)	(14.980)	(18.888)	(21.299)	(13.521)	(16.091)	(16.816)
	Owns land	0.607	0.586	0.555	0.524	0.526	0.507	0.342	0.414	0.461
		(0.488)	(0.493)	(0.497)	(0.500)	(0.499)	(0.500)	(0.475)	(0.493)	(0.499)
	Scheduled caste	0.207	0.251	0.278	0.112	0.168	0.225	0.065	0.092	0.156
		(0.405)	(0.433)	(0.448)	(0.316)	(0.374)	(0.418)	(0.247)	(0.288)	(0.363)
	Scheduled tribe	0.130	0.152	0.165	0.044	0.057	0.072	0.018	0.023	0.036
		(0.336)	(0.359)	(0.371)	(0.205)	(0.232)	(0.259)	(0.134)	(0.150)	(0.187)
	Number of quarters	51,585	146,408	54,710	17,966	47,790	22,567	13,668	42,544	23,013
	Number of women	6,641	18,219	7,544	2,143	5,498	2,735	1,313	3,890	2,423

Note. Means without parentheses and standard deviation in parentheses. Interactions between variables, baseline hazard dummies and squares not shown.

for the third and fourth spell, where around seventy percent of the observations are censored. Generally, censoring increases with parity and time period.[22] This reflects a combination of factors: The timing of the surveys relative to the periods of interest, later beginning of childbearing, falling fertility and finally censoring from the use of sex selective abortions. Even if the censoring is not only due to sex selection, it does underscore the importance of controlling for censoring of the observations and understanding the progress from one birth to the next.

The descriptive statistics also provide a first indication of how the sex ratio at birth changes over time and by spell. For the first spell the sex ratio is very close to the natural for all education group and all three time periods. As an example, among the highly education group for the 1995–2006 period, 51.3 percent of the children born were boys.[23] For the second spell, all but the highly educated group in the last two periods have sex ratios in line with the natural sex ratio. Women with eight or more years of education have 53.1 and 54.3 percent boys in the 1985–1994 and 1995–2006 periods, respectively. This pattern repeats itself for the third spell, except the percentage boys is higher for the high education group (55.3 and 55.9 for the last two time periods). Finally, for the fourth spell the high education group had sixty percent boys in the last period, i.e. after the PNDT act was introduced. Note, however, that for the fourth spell the number of births is substantially smaller and censoring even more important than for the other spells.

India's population has become progressively more urban and this is reflected in the sample. For the first period, 32 percent of the women entering the first spell lived in urban areas. This increases to 35 percent for the second period and to 42 percent for the final period. In addition, the population is also substantially better educated. Women with no education constituted almost sixty percent in the first period but less than thirty percent in the last period. Correspondingly, women with more than eight years of education were just over twenty percent in the first period but almost half in the last period. Part of the increase in education is correlated with the increase in urbanization but clearly the proportion of better-educated women has increased substantially in the

[22] The exception is for women without education where twelve percent of the women who married in the 1972 to 1984 period did not have a first birth at the time of the survey.

[23] There still appears to be some recall error for the group of women without education for the 1972-1984 period, where 52.3 percent of the children born were boys.

rural areas as well. Among the high education group almost seventy percent lived in urban areas during the first period but this had fallen to less than sixty percent in the last period.

The increased in urbanization and education is likely to exert downward pressure on fertility and the high censoring rates for the later periods are evidence of this. Figure 3 provides an indication of just how strong the decline in fertility has been. It shows the average number of children born to women before they turned 35 by their year of birth. Women born in the early 1940s had on average close to five children when they reached 35, but women born in the early 1970s had only just over three children. The low number of children is especially remarkable because it combines all education levels and all area of residence. Hence, fertility in cities is likely substantially lower.

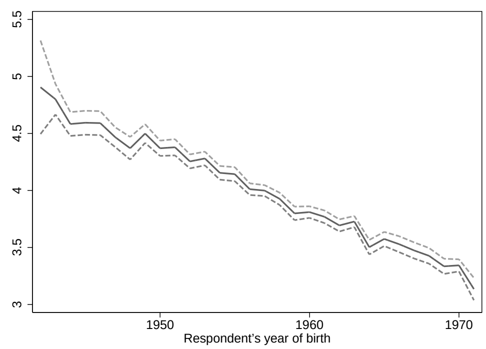

Figure 3: Children ever born at age 35

7 Results

Results are presented using graphs of predictions of the percentage boys born by quarter, the associated 95 percent confidence interval, and the survival curve for a "representative woman" using the method detailed in Section 5. The "representative woman" is a set of values of the explanatory variables chosen to be close to the mean for the relevant sample based on the descriptive statistics provided above. Common to all the "representative women" is that they do not own land and do not belong to either a scheduled caste or a scheduled tribe.

The survival functions show the predicted probability of not having had a birth yet by the length of the spell. The survival curves are included for two reasons. First, they show the predicted progression to the next birth by quarter and how quickly it occur, if at all. Secondly, they provide "weighting" for the associated predicted percentage boys born. The steeper the survival function, the more weight should be assigned to a given spell period because it is based on more births, whereas if the survival curve is flat the period should be given little weight because the predicted percentage boys is based on few births.

7.1 Spell from Marriage to First Birth

The analysis of the first spell is presented for all three education groups for three reasons. First, previous research claim that the largest number of missing girls is for first order births (Jha et al. 2006). Secondly, there are substantial more first births than subsequent births, allowing for a precise estimation of the "natural" percentage boys born in India. Finally, the results provide an indication of whether the length of the first spell is a good indicator for fecundity.

Figures 4 to 6 show the predicted percentage boys born by quarter from marriage to first birth and the associated survival functions for the three education groups for representative women.[24] For the first spell the representative woman is 16 years old at the beginning of the spell for the no education group, 17 years old for the middle education group and 20 years old for the high education group. Each column represents a time period with the top panel showing urban results and the bottom panel rural results. The graphs also show the expected natural rate of boys, approximately 51.2 percent. For comparison, if 55 percent of children born in a given quarter were boys, approximately 14 percent of the female fetuses were aborted. The corresponding numbers for 60 percent and 65 percent boys are approximately 30 percent and 43 percent of the female fetuses aborted.

The most interesting result is how close to the natural sex ratio the predicted percentage boys is for each group and for each period. This is most clearly seen for rural women without education before 1985, who also represent the biggest group (Figure 4d). For these women the predicted sex ratio is aligned almost perfectly with the expected sex ratio. For the other groups there is more volatility in the predicted percentage boys, but nowhere is it statistically significantly larger

[24] The underlying hazards figures for these and all subsequent hazard figures are available on request.

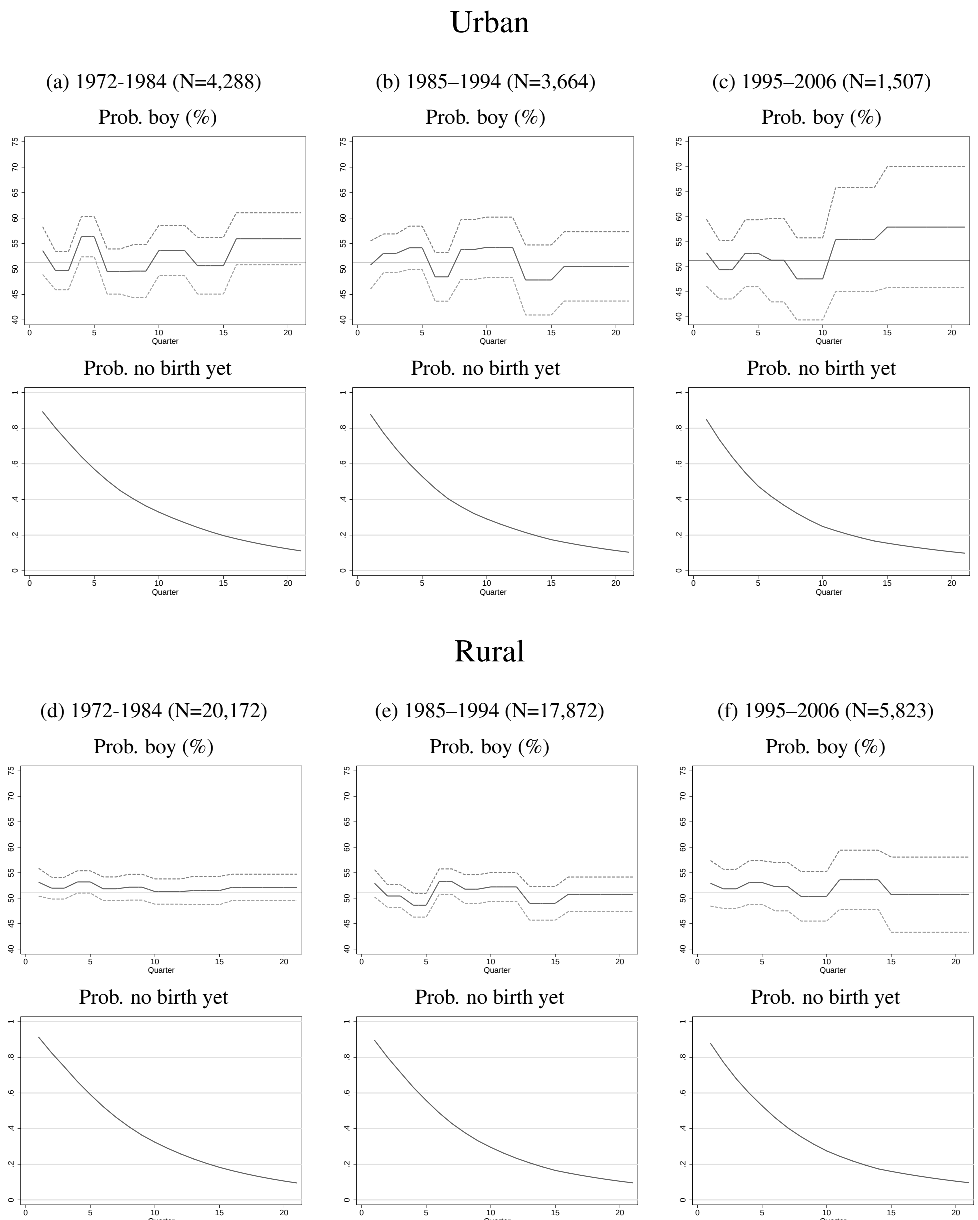

Figure 4: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with no education by quarter. Predictions based on age 16 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

than 51.2 percent.[25] Furthermore, for quarters with more substantial deviations from the natural

[25] The urban no education group for the 1972-1984 period show two quarters where the predicted percentage boys is just statistically significantly higher than 51.2 percent, but this is likely due to recall error not perfectly caught by the method above and the periods around those two quarters are below the natural percentage boys.

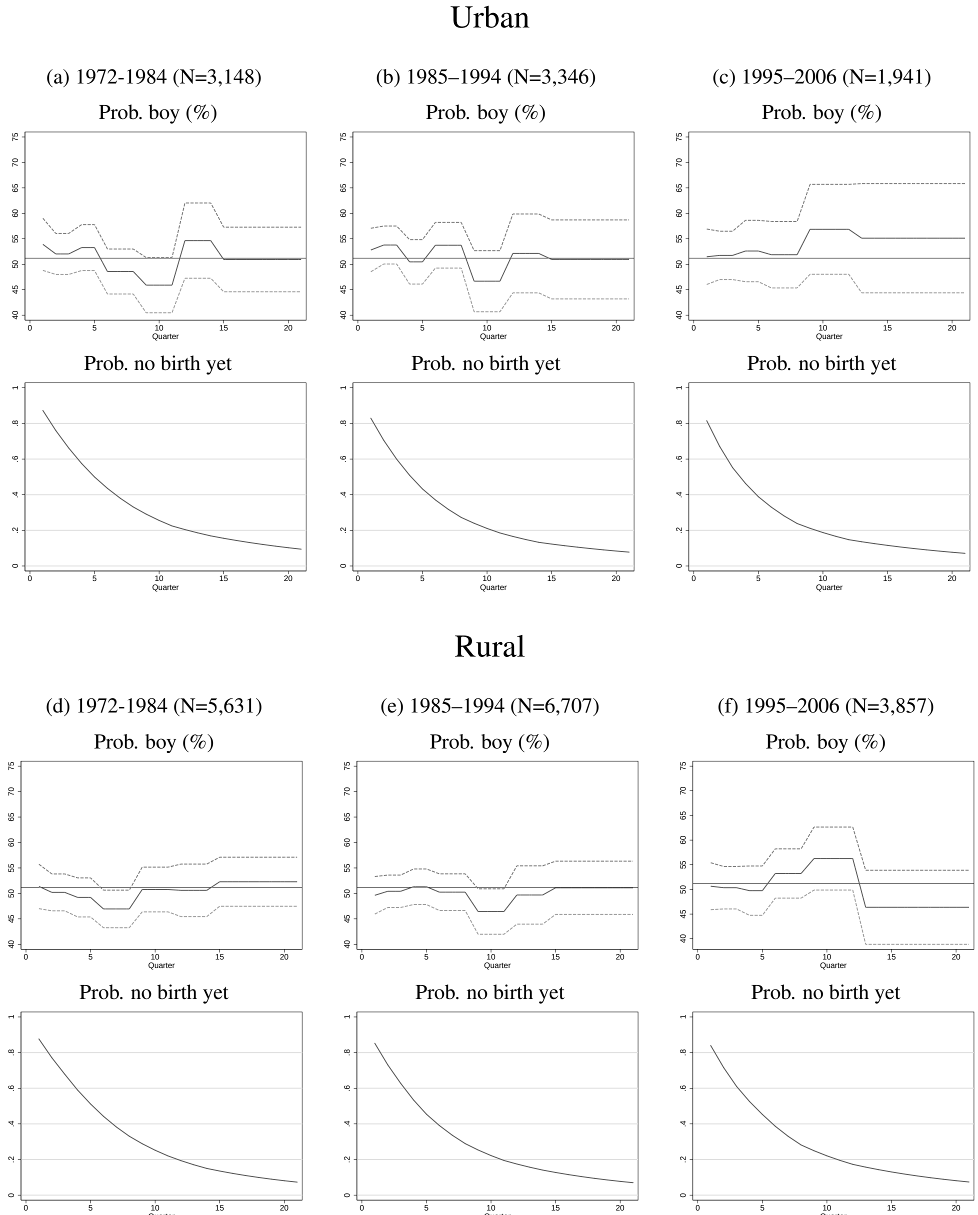

Figure 5: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with 1 to 7 years of education by quarter. Predictions based on age 17 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

sex ratio, the predictions are generally based on fewer births. In other words, it appears that the probability of having a boy is exactly the same in India as it is in other places.

For the group most likely to use sex selection, highly educated, urban women in the 1995–2006

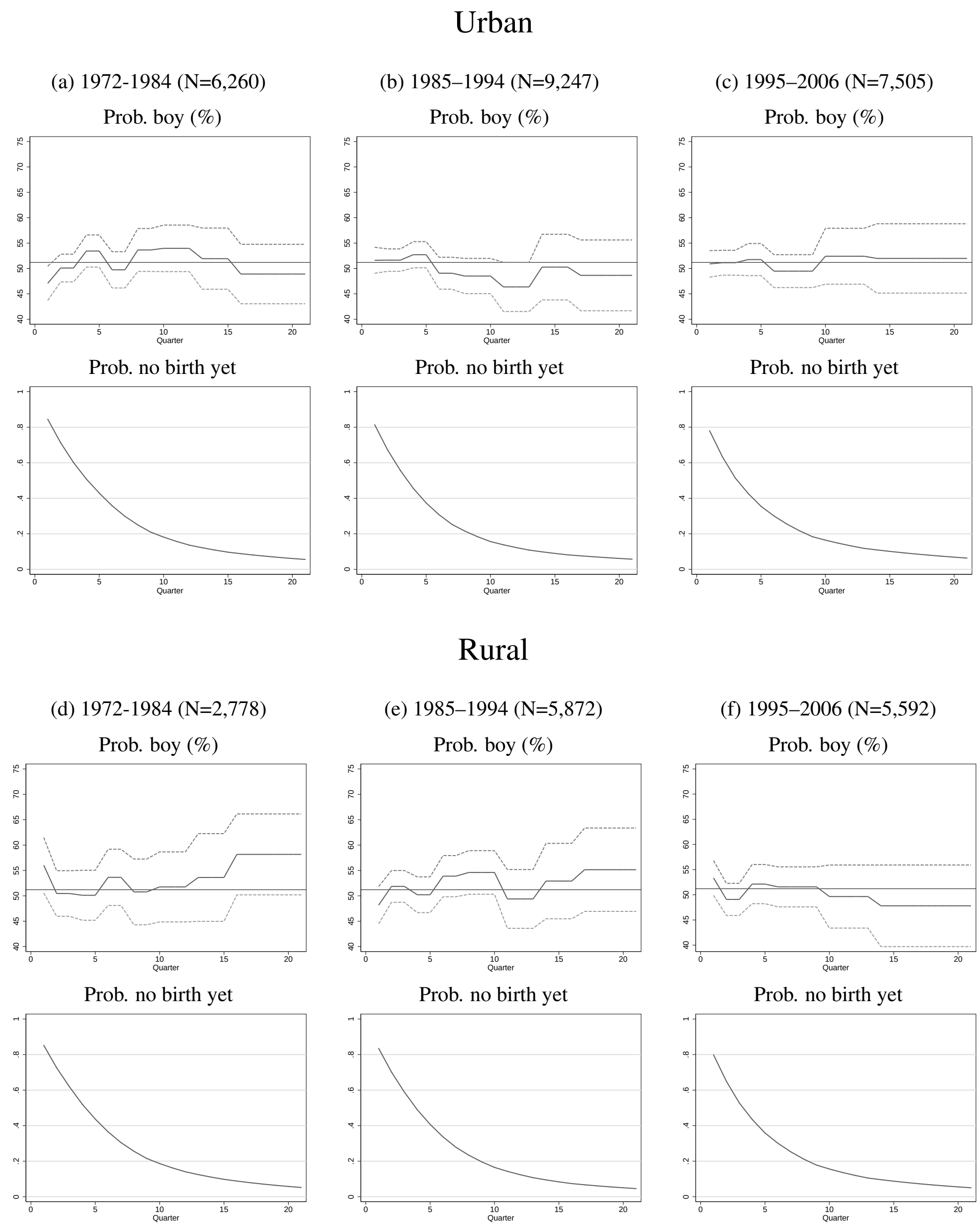

Figure 6: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with 8 or more years of education by quarter. Predictions based on age 20 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

period, the predicted percentage boys is also almost perfectly aligned with the expected percentage boys (Figure 6c). Hence, there is no evidence that Hindus in India use sex selection on first births. This cast serious doubts on the data used by Jha et al. (2006) and their results.

For all education group and for all periods more than ninety percent of women will have had their first child within 21 quarters of being married and the proportion is increasing in education. In addition, 70 to 85 percent of women will have had their first child within 2.5 years (ten quarters) of their marriage and the space between marriage and first birth has become shorter over time. The most likely explanation for the reductions in spell lengths and the increase in the number of women who have their first child before 21 quarters is improvements in health status. This is consistent with the differences between education groups where more educated women are also more healthy and therefore more likely to conceive. There are two implications of this. First, it reinforces the need for estimating the models separately for different education levels. Secondly, it confirms that the length of the first spell can serve as a suitable proxy for fecundity and that Hindu women in India have their first birth very soon after marriage, even among highly educated, urban women.

7.2 High Education Group

If lower fertility is the driving force behind sex selection, women with the most education should be the earliest adopters and use it for lower parity children than those with less education. The predicted percentage boys for the second, third and fourth spells for women with eight or more years of education show that this is, indeed, the case. As for the first spell, the results are presented using a representative woman with the same characteristics as above, except that the ages at the beginning of the spell are 22, 24 and 25 years for the second, third and fourth spell and that the first spell length is set equal to 16 months.

Figures 7 and 8 show the second spell predicted percentage boys and the survival curves by spell length for urban and rural women.[26] For both figures, the top panel shows the results if the first child was a girl and the bottom panel the results if the first child was a boy. Not surprisingly, there is no evidence of sex selection for the 1972-1984 period; the predicted percentage boys match up closely with the natural. There is some evidence that spacing is shorter after the birth of a girl than after a boy, but the differences between the survival curves in the beginning of the spells are

[26] Recall, zero quarters is equivalent to nine months after the first birth.

not large. The largest difference at five quarters is for urban women and that is only around five percentage points. This is consistent with more educated women being more aware of the potential negative effects of close spacing.

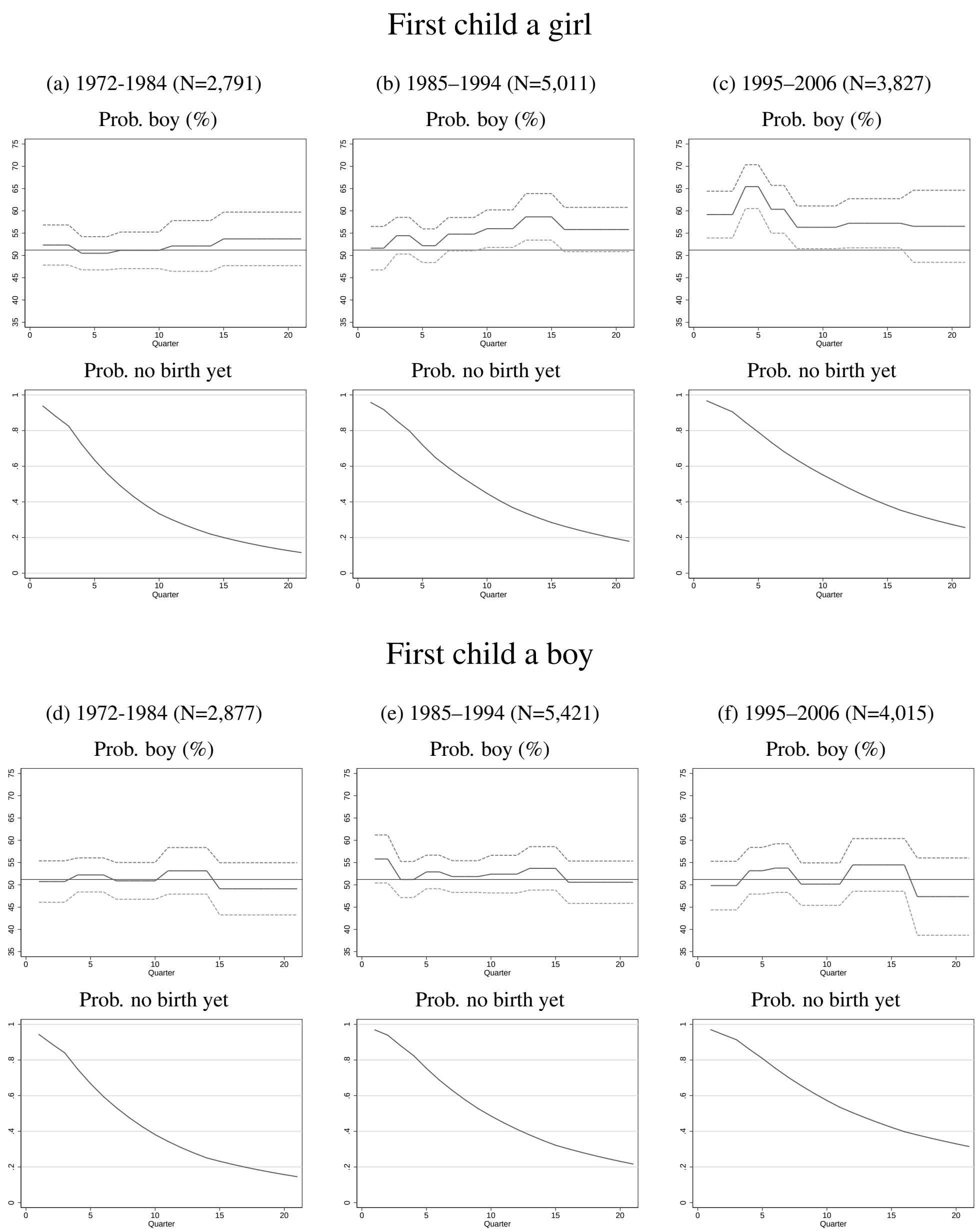

Figure 7: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with 8 or more years of education by quarter. Predictions based on age 22 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

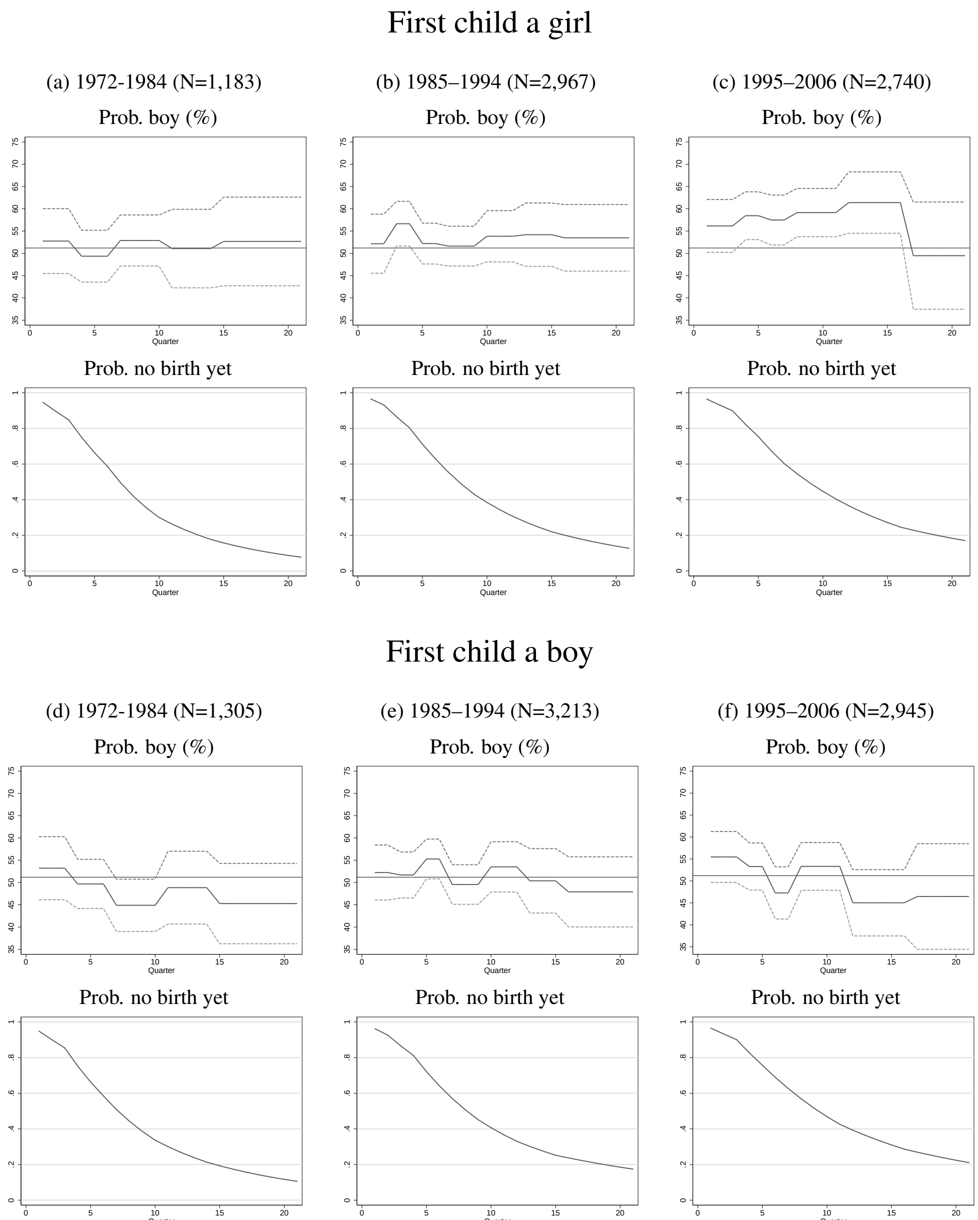

Figure 8: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with 8 or more years of education by quarter. Predictions based on age 22 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

The first substantive evidence of sex selection is for urban women with one girl in the 1985–1994 period (Figure 7b). The percentage boys begins at the natural level but then increases to almost 60 percent after which it drops slightly. The use of sex selective abortions becomes even more apparent for the 1995–2006 period, where the percentage boys born to highly educated women with one girl begins at just below 60 percent, increases to above 65 percent, followed by a decline to just over 55 percent. These results are especially remarkable given that the numbers are for all of India, not just the states with traditionally strong son preferences.

Of particularly interest is how the predicted percentage boys changes with spell length. The development in the percentage boys by quarter for the 1995–2006 period (Figure 7c), is consistent with some women changing their decision to use sex selection after going through one or more abortions. The results are for the median level of fecundity, and it is therefore unlikely that the decline is due to lower fecundity for women who give birth later in the spell.[27] The 1985–1994 period (Figure 7b) shows the same fall at the end of the observed spell, but it also begins at the natural level, consistent with increasing access and acceptance of sex selection over the early period. Women who began their second spell early in the 1985–1994 period were less likely to have access than women who began their second spell later in the period. The 1985–1994 results can be thought of as the pattern that will prevail as sex selection is introduced, and the 1995–2006 results show the pattern with a mature and readily available technology. That sex selection for the 1995–2006 period is used more early in the spell compared to later also shows that not taking account of timing of births and the potential censoring of birth spell can lead to an upward bias in the final sex ratio of second born children.

For rural areas there is no sign of sex selective abortion during the 1985–1994 period, but for the 1995–2006 period the percentage boys born to highly educated women who have a girl as their first child increases gradually from 55 percent to over 60 percent until dropping off to the normal percentage boys (Figure 8c). Hence, there appears to be easy access to pre-natal sex determination even in rural areas. The difference between urban and rural women for the 1985–1994 period is more likely to the result of differences in fertility than in access to pre-natal sex determinantion as shown below. As more and more rural women have only two children the use of sex selection goes up. The implication is that rural areas lag behind urban areas but not by much and that the

[27] It is, of course, possible that women find it harder to conceive during the second spell, but the survival curve would be very close to the survival curve for the first period if the female fetuses were carried to term.

expected pattern for the next round of the NFHS for rural educated women should closely mirror what is found for the 1995–2006 period for urban women.

Women who had a boy as their first child do not appear to be using sex selective abortions. This holds for urban and rural areas and for all time periods. In addition, women with a boy are less likely to have a second birth within the 21 quarters. This is especially prevalent for urban women, although for rural women show the same pattern. For urban women with one boy, more than thirty percent are not predicted to have their second child within the 21 quarters covered here. This means that six years have passed since the birth of their first child and given the flatness of the survival function it it unlikely that many of these women will ever have a second child. Even for rural women around twenty percent of women will likely not have a second child.

The result for the third spell are shown in Figures 9 and 10 for the representative urban and rural women. The top panel shows the result if the first two children were girls, the middle panel if the couple had one boy and one girl, and the bottom panel if the first two children were boys. Again there is no evidence of sex selective abortions before 1985, but clearly there is more volatility in the predicted percentage boys due to the lower sample sizes.

In urban areas there is clear evidence of sex selective abortions among women with two girls for both 1985–1994 and 1995–2006. For the first of these periods, births early in the spell are close to the natural sex ratio, but births after the fifth quarter are significantly above the natural level at around 65 percent boys.[28] Interestingly, there does not appear to be an increase in the use of sex selective abortions in the 1995–2006 period compared to the 1985–1994 period. This can be explained by the increase in the use of sex selective abortions for second births combined with lower desired fertility. Fewer women have two girls and given the apparent ready access to pre-natal sex determination those who do are more likely to do so by choice.[29] Even then, the percentage boys born is between 60 and 65 percent. This pattern of sex selection for women with two girls is mirrored by rural women. As expected, given that most educated women do not want

[28] The sex ratio increases substantially for births after the 15th quarter, but given the low number of births it is not clear how much one can learn from this increase.

[29] For urban women, 25 percent of the sample in the first period had two girls, while only 22.7 percent of the third period sample had two girls.

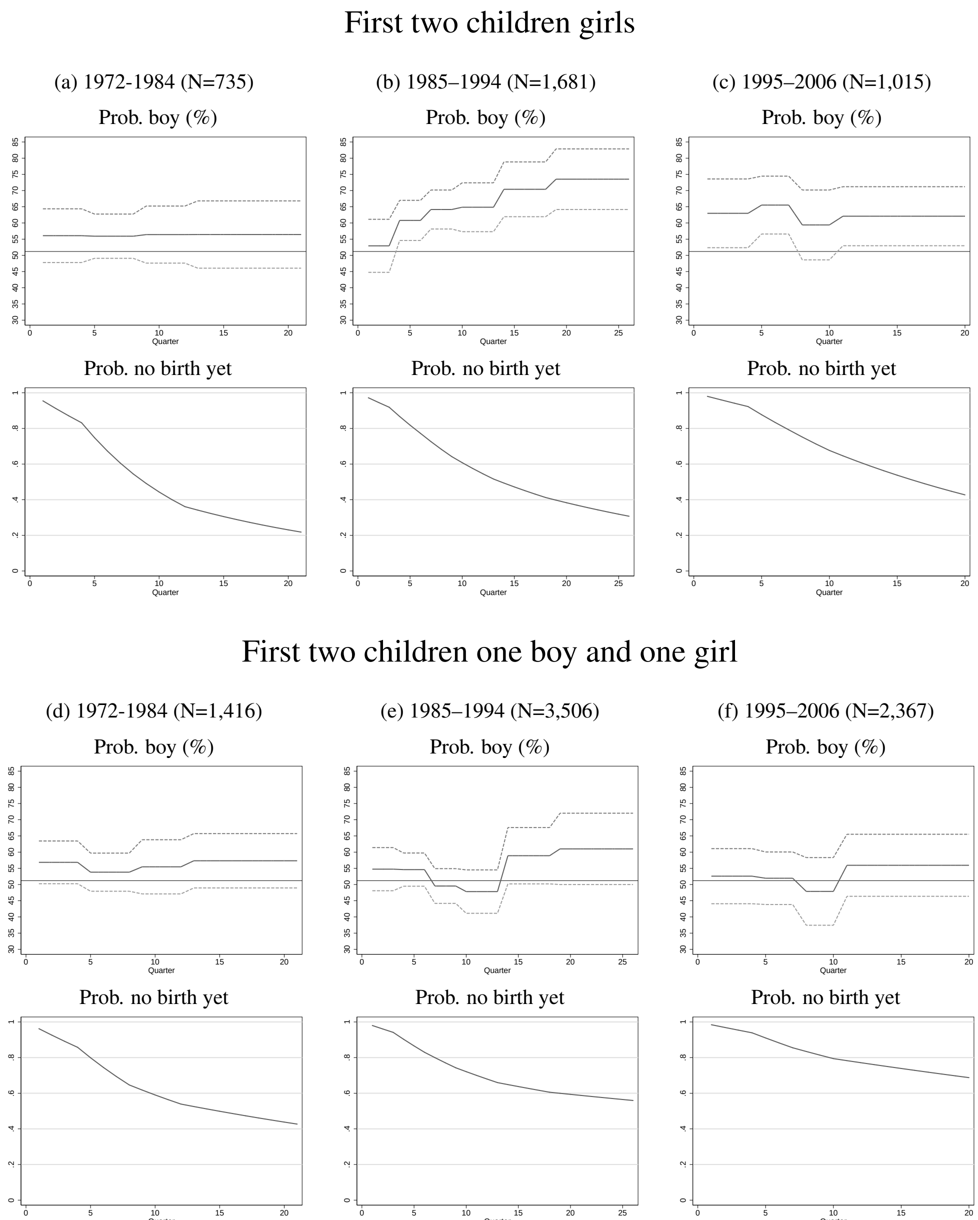

Figure 9: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter. Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

more than two children, there is no evidence for declining use of sex selection toward the end of the spell.

Women with either two boys or one boy and one girl as their first two children do not appear

First two children boys

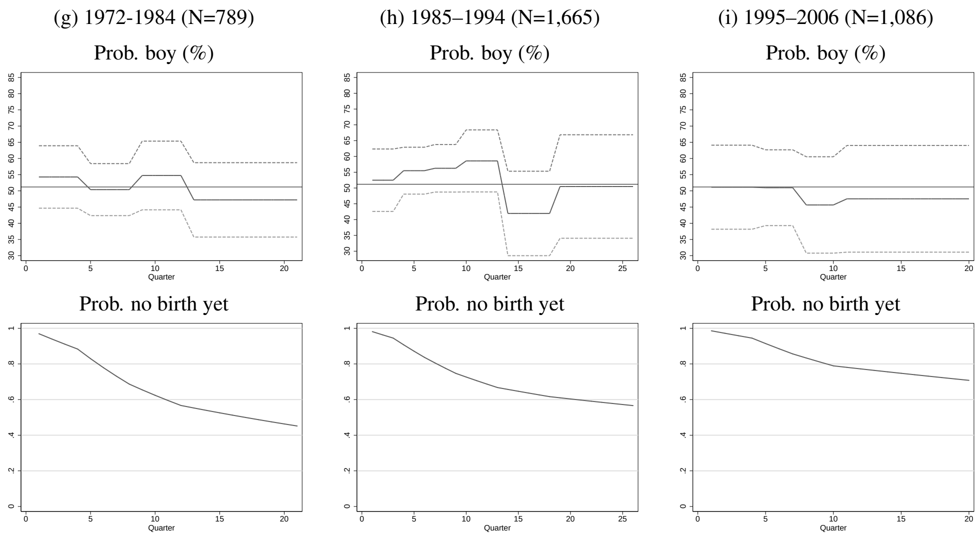

Figure 9: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter. Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

to use sex selection. Relatively few of these women go on to have a third child and the proportion that does declines substantially over time. This is best illustrated by the urban women in the 1995–2006 period (Figures 9c, f, i). Close to sixty percent of the women with two girls will have a third child, while only around thirty percent of those with at least one boy will have a third child. For comparison, during the 1972-1984 period eighty percent of the urban women with two girls had a third child and almost sixty percent of those with at least one boy had a third child. The numbers are higher for rural women, but the pattern is the same.

Relative few women with eight or more years of education have three children and even fewer have four. This means that caution should be used in interpreting the results and in the interest of space the graphs are not presented here.[30] As for the other spells there is no evidence of sex selection for the first period. Again, families without boys are the main users of sex selective abortions. For urban women with three girls the predicted percentage boys is between 60 and 75

[30] Graphs and underlying estimation results are available on request.

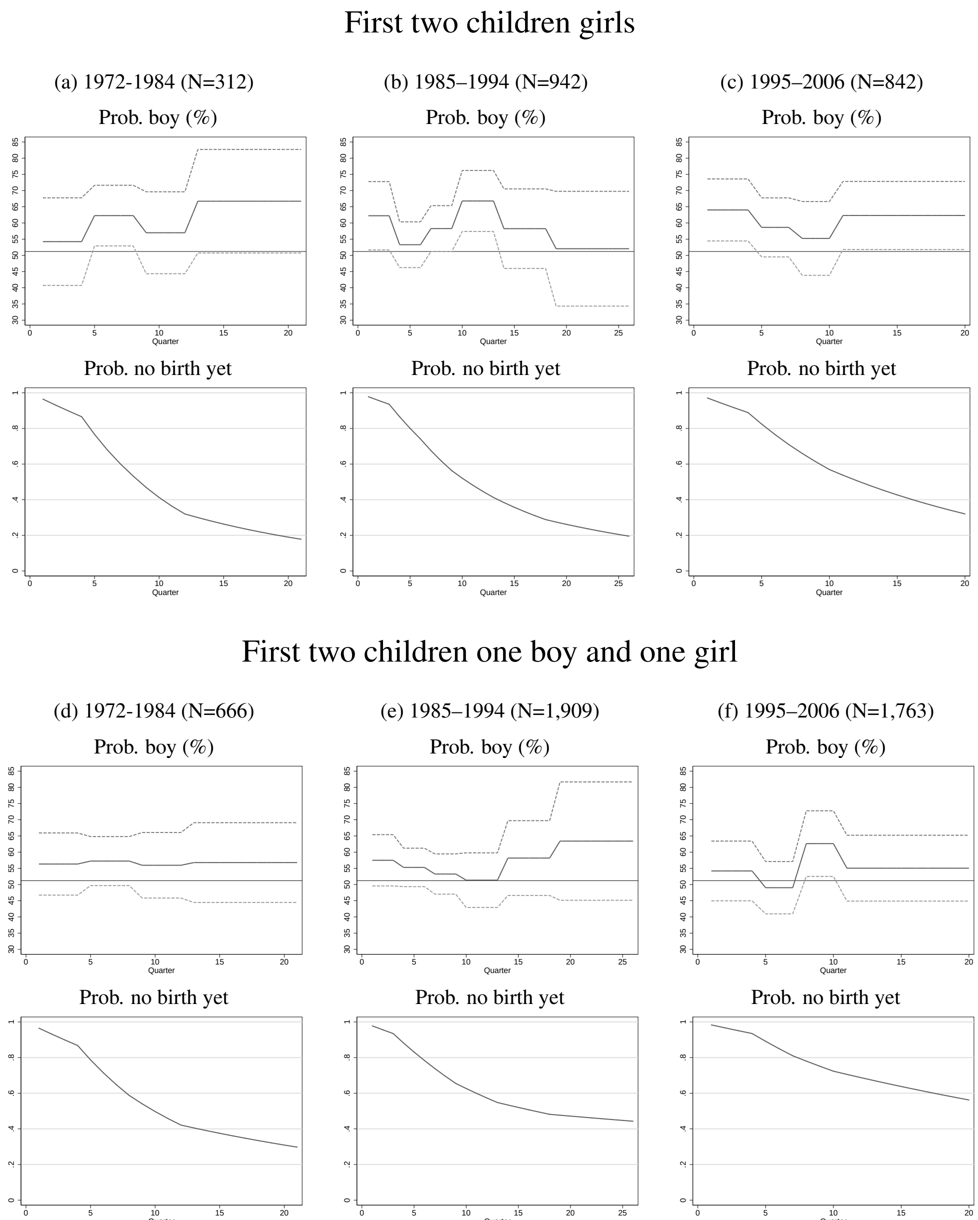

Figure 10: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter. Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

for both the 1985–1994 and 1995–2006 time periods. The numbers for rural women with three girls are in line with those for urban women. It is worth noting that there are only 211 and 155 observed births for urban women for the two time periods, and only 167 and 178 births for the rural

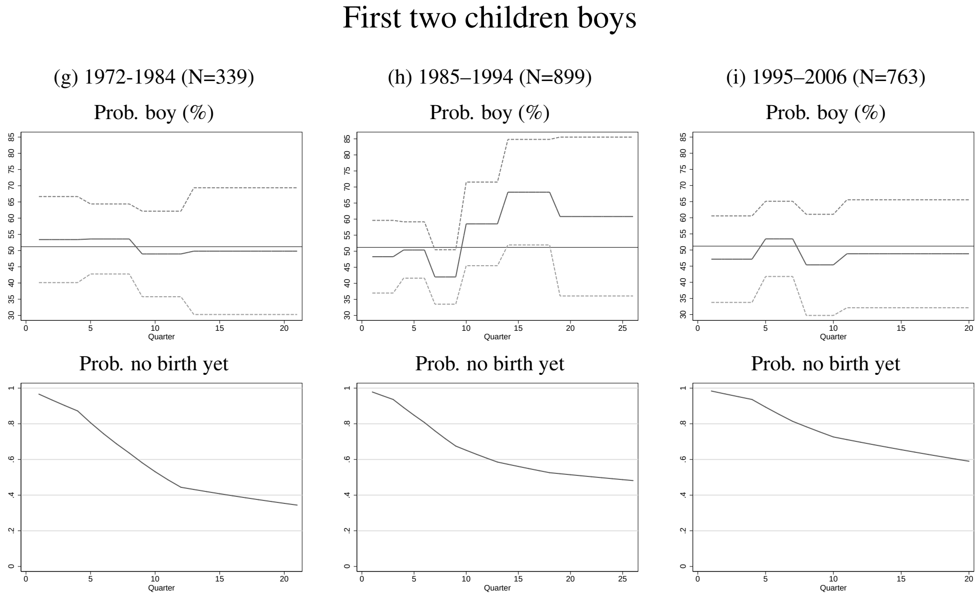

Figure 10: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter. Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

women. Urban women with with two girls and one boy show a predicted percentage boys above 65 percent for all quarters, but this result is based on only 292 births. Finally, it is interesting that the probability of having a fourth child is higher for women with three boys than it is for women with two boys and one girl. This might be evidence of a slight gender balance preference, although there is absolutely no evidence that women are aborting boys and the number of observed births to both groups is very small.

7.3 Desired Fertility and Sex Selection

To further examine the connection between fertility and sex selection Figure 11 presents results by desired fertility for the second spell for women with 8 or more years of education whose first child was a girl. A majority of women want two or fewer children and that number is increasing over time. To achieve a large enough sample all second spells that began in the period 1985 to 2006 are therefore used. This allows for non-proportional hazards by area of residence, sex of previous

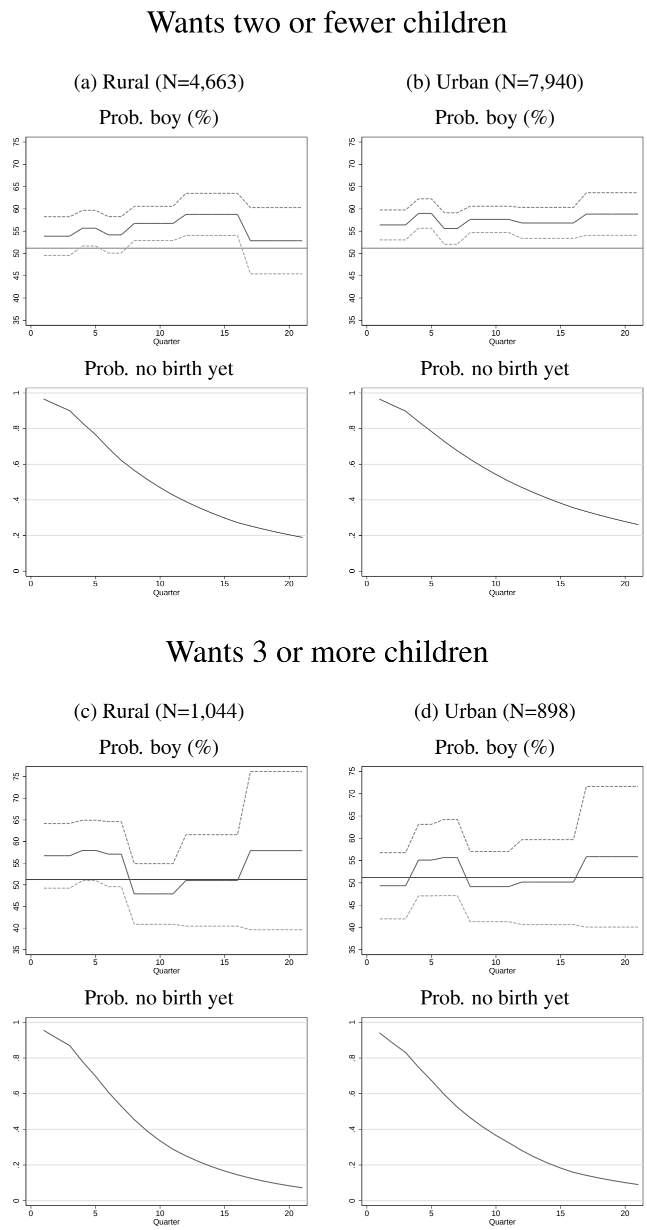

Figure 11: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for women with 8 or more years of education by quarter. Predictions based on first child a girl and woman age 22 at first birth. N indicates the number of women in the relevant group in the underlying samples.

child and desired fertility to be estimated. For both rural and urban the left figure shows the results for women who want two or fewer children and the right shows the results for women who want three or more children.[31]

The proportion of women who have a second child is substantially higher among women who want three or more children than among those who want two or fewer. There is little or no evidence of sex selection for women with higher desired fertility (Figures 11c, d), whereas sex selection is clearly used by women who want two or fewer children, especially among urban women (Figures 11a, b). For urban women who want only two children there is no change in the use of sex selection with the length of the spell as expected for women who do not want more than two children.

7.4 Middle and Low Education Groups

For women with one to seven years of education, a representative woman is used with the first spell length set to 16 months and the ages at the beginning of the spell 19, 21 and 24 for the second, third and fourth spells, respectively. There are two factors that make it difficult to establish whether sex selection is used among the middle education group. First, there are fewer observations for this group compared with the high education group. That problem becomes especially acute when looking at the higher spells divided by sex composition of previous children. Secondly, even for those spells with more information there is substantially more noise in the 1972-1984 results than for the high education group.[32]

For the second spell there is no evidence of sex selective abortions for either urban or rural women and the proportion of women who have a second birth remains high across all three time periods. The same is the case for the third spell, although there signs of falling fertility in that the proportion of women who have a third child is decreasing, especially among those with one or two boys.[33] Even for the fourth spell, women with three girls does not show statistically significant use of sex selective abortions and this holds for both urban and rural women. For both urban and rural

[31] A concern with using stated desired fertility is that it is potentially endogenous to the number of children a woman has actually had. These results are therefore best treated as suggestive. See Rasul (2008) and Ashraf, Field and Lee (2009) for discussions of stated fertility preferences and their effect on fertility.

[32] The results are available on request.

[33] There are relatively few urban women with two girls in this education group and although there might be some evidence of sex selective abortions the results are too noisy to conclude anything with any degree of confidence.

women there has been a substantial decline in the number of women who have a fourth child and again this decline is concentrated among women with one or more boys.

For women with no education, the starting ages are 18, 21 and 23 for the second, third and fourth spells, respectively. Although there are not a large number of women in urban areas without education there are substantial numbers in the rural areas.[34] For urban women without education there is little evidence of sex selection for any of the spells.[35] Although the percentage boys for rural women with one girl for the second spell is a very precisely estimated flat line for the 1985–1994 period, the percentage boys for the third period is between 55 and 60 from the seventh to the thirteenth quarter and this percentage is statistically significant.[36] Two things makes this result puzzling. First, there is little evidence of women stopping childbearing once they have a son. Even among rural women with two sons almost ninety percent go on to have a third child and eighty percent of women with three boys have a fourth child. Hence, fertility remains very high for this group and therefore also the probability of having at least one son without the use of sex selective abortions. Secondly, one would expect more use of sex selection for women with two girls in the third spell or three girls in the fourth spell, but for both the predicted percentage boys is close to the natural rate.

7.5 Comparing the Hazard and Standard Models

Two outcomes are of special interest for most studies of sex selective abortions: the sex ratio and the number of abortions. We therefore now turn to the differences in predicted number of abortions and sex ratio between the standard logit model approach and the hazard model suggested here.

A major difference between the two methods is how censored observations, that is women who do not have a next parity child, are treated. The logit model uses only observed births and the predicted sex ratio and percentage boys are simply what we observe for the relevant uncensored sample. Hence, we do not even need to estimate the logit model if we are interested only in the

[34] The results are available on request.
[35] Urban women with three girls do have a higher percentage boys, but only 208 are in this group.
[36] These quarters account for more than twenty percent of the births or slightly more than 700 births.

uncensored sample, which is what prior research has focused on.[37] The hazard model instead uses the entire sample, including both censored and uncensored observations, for predictions. This allows us to predict not only how many women will have the next birth but also what the final sex ratio will for that parity.

Two factors combine to determine the differences in the predicted sex ratio and the number of abortions between the two models. First, how many women have censored spells but will go on to have the next birth. Second, to what extent use of sex selection changes within a spell. If there are few censored spells or the women with the censored spells have low likelihoods of having a next birth the two models produce similar predicted sex ratios and number of abortions. If there is little change in the use of sex selective abortions over a spell then the predicted sex ratio is going to be similar for the two models, although the predicted number of abortions will differ depending on how many censored observations there are and how likely they are to exit with a birth. The problem with the logit model is that does not provide precise information on final sex ratio and abortions when the use of sex selective abortions varies with the timing of births. If the percentage of boys is increasing in spell length the number of abortions and the sex ratio predicted by the logit model is going to be too low. If the percentage of boys instead is decreasing as the spell becomes longer the logit model will overestimate the sex ratio. These problems are exacerbated the more censored women there who have not completed child bearing.

To determine the predicted number of abortions we make use of the normal sex ratio. We expect 105 boys to be born per 100 girls born, which corresponds exactly with what we observe for first births in India as seen above. Hence, if we observe b boys, then we should expect to see $b\frac{100}{105}$ girls. If we observe g girls the number of abortions is therefore $b\frac{100}{105} - g$. The standard approach is to use the observed sex ratio for a given parity and calculate the number of abortions using this formula. For the hazard model it is, however, possible to calculate the number of predicted abortions for each time period based on the predicted sex ratio and the number of women predicted to exit with a birth in that period using the same formula. Summing across the entire period covered by the

[37] The predicted outcomes for the censored samples using the Logit model is available upon request. These predictions are done under the assumption that all censored women will eventually give birth.

hazard model provides a predicted total number of abortions.[38]

We focus on those cases where there is evidence of sex selective abortions for the spells from parity one to two and two to three for women with eight or more years of education.[39] We begin with women who had a girl as their first child. Table 4 shows the number of women who enter the second spell, the number of births by the end of the six year period following the first birth, observed in the case of the logit model and predicted in the case of the hazard model, predicted number of abortions, and the sex ratio.

Table 4: Comparison of Hazard and Logit Models for Second Birth for Women with 8 or More Years of Education and 1 Girl

	Urban		Rural	
	Logit	Hazard	Logit	Hazard
	1985–1994			
Starting population[a]	5011	5011	2967	2967
Second birth[b]	3725	4059	2427	2665
Number of abortions	250	282	90	100
Sex ratio (boys per 100 girls)	120.5	121.1	113.3	113.4
	1995–2006			
Starting population[a]	3827	3827	2740	2740
Second birth[b]	1957	2763	1659	2328
Number of abortions	296	393	176	252
Sex ratio (boys per 100 girls)	143.7	140.9	130.7	131.3

Note. The logit models are estimated using only uncensored observations. The hazard models are estimated using all observations and covers the period from beginning of spell to 6 years (24 quarters) after the birth of the first child.
[a] Number of women in period who had a girl as their first child.
[b] For logit model the number of women who are observed to have a second birth. For hazard model the predicted number of births that will occur between beginning of the spell and 6 years (24 quarters) after the birth of the first child.

The table nicely illustrates the different cases. For rural women in the 1985–1994 period relative few censored women are predicted to go on to have a second birth (only eight percent of the starting population is predicted to have a second birth and have not yet had one). In addition,

[38] There is no difference between the two method in the predicted number of abortions for those women for whom we observe the subsequent birth.

[39] The results for the spell from parity three to four are available on request. They are not presented because the number of educated women who end up with three births is relatively small and the probability of having a fourth birth is small, which makes comparing the logit and hazard models with any precision difficult.

there is only some evidence of sex selective abortions. Hence, the predicted sex ratios for the logit and hazard models are close. There are slightly more abortions for the hazard model, but the abortion rates (number of abortions per birth) is fairly similar between the observed births and those predicted to have a birth later. Urban women in the same time period follow a similar theme with relative few censored women predicted to move on to a second birth (around seven percent of starting population), but the abortion rate for those additional births is 9.6 abortions per 100 births whereas it is only 6.7 for the observed births, as we would expect from the increasing probability of having a son with increasing spell length seen in panel (b) in Figure 7. Despite this the resulting difference in observed versus final sex ratio is relatively small because of the small number of censored women predicted to exit with a birth.

Not surprisingly there are substantial more censored observations for the 1995–2006 period. For urban women about 21 percent of the starting population are predicted to have a second birth but have not yet had one and for rural women it is 24 percent. This allows for potentially larger differences in predicted sex ratios and number of abortions. Despite the large number of censored observations for rural women there is relatively little difference in the predicted sex ratio with the logit model predicting 130.7 boys and the hazard model 131.3 per 100 girls as we would expect with the relatively stable predicted probability of having a boy over the spell shown in Figure 8, panel (c). The more interesting difference is for urban women in the 1995–2006 period. The hazard model predicts almost a 100 more abortions by the end of the spell but shows that the logit model is likely to *overestimate* the final sex ratio with the logit model predicting 143.7 boys and the hazard model 140.9 boys per 100 girls. This is consistent with the very high probability of having a son early on in the spell followed by a somewhat lower probability for later births seen in panel (c) in Figure 7.

Table 5 presents the predicted outcomes for urban and rural women for the third spell split by the sex composition of their first two children and when the spell began.[40] For urban women who had one boy and one girl as their first two children there is little difference in the predicted sex

[40] There is no evidence of sex selective abortions and very few births for women with two boys and those predictions are therefore not presented.

Table 5: Comparison of Hazard and Logit Models for Third Birth for Women with 8 or More Years of Education

	Urban		Rural	
	Logit	Hazard	Logit	Hazard
	1985–1994			
	1 boy and 1 girl			
Starting population[a]	3506	3506	1909	1909
Third birth[b]	1397	1653	1057	1279
Number of abortions	34	47	62	79
Sex ratio (boys per 100 girls)	110.4	111.3	118.4	119.2
	2 girls			
Starting population[c]	1681	1681	942	942
Has third birth[b]	1005	1149	719	820
Number of abortions	227	275	89	102
Sex ratio (boys per 100 girls)	168.7	173.8	135.7	135.6
	1995–2006			
	1 boy and 1 girl			
Starting population[a]	2367	2367	1763	1763
Third birth[d]	502	740	543	887
Number of abortions	10	18	43	79
Sex ratio (boys per 100 girls)	109.2	110.3	123.5	126.2
	2 girls			
Starting population[c]	1015	1015	842	842
Third birth[d]	396	566	430	607
Number of abortions	86	121	87	125
Sex ratio (boys per 100 girls)	165.8	164.5	160.6	161.6

Note. The logit models are estimated using only uncensored observations. The hazard models are estimated using all observations and covers the period from beginning of spell to 7.25 years (29 quarters) after the birth of the second child for the period 1985–1994 and from beginning of spell to 5.75 years (23 quarters) after the birth of the second child for the period 1995–2006.

[a] Number of women in period who had one boy and one girl as their first two children.

[b] For logit model the number of women who are observed to have a third birth. For hazard model the predicted number of births that will occur between beginning of the spell and 7.25 years (29 quarters) after the birth of the second child.

[c] Number of women in period who had two girls as their first two children.

[d] For logit model the number of women who are observed to have a third birth. For hazard model the predicted number of births that will occur between beginning of the spell and 5.75 years (23 quarters) after the birth of the second child.

ratio between the logit and hazard models for either period and although higher than the natural sex ratio it is not by much. As further evidence that fertility is declining dramatically the predicted likelihood for this group of exiting the third spell with a birth has fallen from 47 percent to 31 percent.

Probably the most dramatic example of the difference between the logit and hazard model is for urban women with two girls in the 1985–1994 period. As figure 9, panel (b) shows the predicted sex ratio increases sharply with spell length. Hence, although there are only 144 predicted more births for the hazard model, the difference in number of abortions is 48 with the hazard model predicting 275 abortions and the logit model 227. This translates to 23.9 abortions per 100 births for the hazard model. The resulting sex ratio is 173.8 for the hazard model versus "only" 168.7 for the logit model. In other words, it appears that the logit model underestimate the spread of the use of sex selective abortions during the expansion of access to pre-natal sex determination for this group of urban women.

Use of sex selection for urban women with two girls appears to actually have fallen from the 1985–1994 period to the 1995–2006 period. This might seem like a good sign but is mainly an indication that sex selective abortions are being use more during the spell to second birth for this group of women for the 1995–2006 period. The result is fewer women with two girls for their first two births and those who do are substantially less likely to have a third birth than they were before.

During the 1985–1994 period there is relatively little difference between the logit and hazard model for rural women. Clearly, rural women had access to pre-natal sex determination and subsequent abortion during this period; for women with two girls as their first two children there are a predicted 12.4 abortions per 100 births for the 1985–1994 period.

There are substantially larger differences between the logit and the hazard model predictions for the 1995–2006 period with the logit model underestimating the final sex ratio no matter the sex composition of the first two births. For women with one boy and one girl the abortion rate for the censored observations is 10.5 per 100 births whereas it is 7.9 for uncensored observations. This combined with only around 60 percent of the predicted births actually observed leads the final sex ratio to be 126.2 for the hazard model compared to 123.5 for the logit model. That the abortion rate and the final sex ratio is as high as it is for women who already had one boy shows that the "funeral pyre" hypothesis does not appear to hold for this group of women.

The logit model also underestimates the sex ratio for women with two girls but the effect is less pronounced with the abortion rates 20.2 and 21.5 per 100 births for uncensored and censored observations. The result is a sex ratio of 160.6 for the logit model and 161.6 for the hazard model. This means that among women with two girls, rural women are close to having the same use of sex selective abortions as urban women and points to the large increase in the use of sex selective abortions between the two periods for rural women. Furthermore, although fertility has decreased for rural women over time 50 percent of of rural women with one boy and one girl and 72 percent of rural women with two girls are predicted to have a third child. Hence, there is little evidence of the use of sex selective abortions slowing down at all.

7.6 Predicted Lifetime Number of Children and Abortions

An important advantages of the new method is that one can estimate both fertility and use of sex selective abortions as women move through their childbearing years. Table 6 shows predicted fertility and number of abortions for the two samples of women with eight or more years of education.[41] The predictions are done as follows. For a given spell, a woman's probability of having a birth in each quarter is calculated and she either has a child or not based on a binomial distribution with this probability. If she has a child she is assigned a boy or a girl, again based on a binomial distribution with the probability of having a son based on the predicted probability from the estimations above. Once she has a child she leaves that spell and her starting age for the next spell is calculated. If she does not have a child before the quarters run out she is assumed to have stopped childbearing and she does not enter the next spell. This is done for all four spells, so fertility is capped at four. Because the samples might differ in composition between periods, the results for the two periods are presented for both women who married in the 1985–1994 period and women who were married in the 1995–2006 period. The presented results are averages of 80 repetitions.

Fertility has fallen over time; predicted completed fertility for urban women went from 2.3 during the 1985–1994 period to just barely over 2 during the 1995–2006 period. Hence, fertility

[41] In other words, the same women who were used to estimate the results for the spell from marriage to first birth above. There are 15,119 women in the 1985–1994 sample and 13,097 women in the 1995–2006 sample. For more information see Table 2.

Table 6: Predicted Lifetime Fertility, Sex Ratio, and
Abortions for Women with 8 or More Years of Education

Sample period	1985–1994		1995–2006	
Estimation Results from	Period 2	Period 3	Period 2	Period 3
	Urban			
Fertility	2.3	2.1	2.3	2.0
Percentage boys	52.7	53.0	52.6	53.2
Abortions per 100 births	3.0	3.6	2.9	3.8
Abortions per 100 women	7.1	7.6	6.7	7.8
	Rural			
Fertility	2.8	2.6	2.7	2.5
Percentage boys	52.5	52.7	52.4	52.7
Abortions per 100 births	2.5	2.8	2.4	2.9
Abortions per 100 women	6.9	7.2	6.6	7.3

Note. Prediction of fertility and abortions are based on the results above. The samples consist of all women who married in a given period. All numbers presented are for women predicted to have one or more births. Results are average values of 80 repetitions of the predictions.

for urban women is likely to be below replacement. Rural women also showed a decline in fertility over time going from 2.8 to 2.5.

For all groups, the number of abortions expected before childbearing is over increased substantially between the two periods. Ignoring the compositional effects and comparing the first and last columns for urban women, the number of abortions per 100 women have gone up by almost 10 percent (7.1 to 7.8) and the number of abortions per 100 births have gone up by over 25 percent (3.0 to 3.8). For rural women, abortions per 100 women increased 5 percent (6.9 to 7.3), while abortions per 100 births increased over 15 percent (2.5 to 2.9).

Of particular note is the high level of abortions for the 1985–1994 period, especially for rural women. Rural women have more children than urban women and therefore need fewer sex selective abortions to ensure a son, but given differences in fertility the numbers of abortions for rural women are not far behind those of urban women. Educated rural women obviously had few constraints on access to pre-natal sex determination even when the techniques were relatively new. As fertility continues to fall, we are likely to see further increases in the use of sex selective abortions.

8 Conclusion

This paper introduces a novel approach that allows for joint estimation of fertility and sex selective abortions using a non-proportional competing risk hazard model. Three hypotheses are examined. First, lower desired number of children is a driving factor in the decision to use sex selective abortion. Only women with low expected fertility show evidence of sex selection. In both urban and rural areas, better-educated women—those with eight or more years of education—use sex selection, and as fertility falls they use it for earlier parities. In addition, women wanting more than two births did not use sex selection on the second birth. Secondly, parents overall appear to have a preference for one son rather than many sons. There is little evidence that women with one or more sons use sex selection, and their probability of a next birth declines substantially once they have a son. The exception is for rural women with one son and one daughter, for whom the use of sex selective abortions have increased over time, although the probability of having a third child has also decreased. Third, the legal steps taken to combat sex selective abortion have not been able to reverse the practice. The use of sex selection has been increasing over time and is higher now than before the PNDT Act was passed. For all of India, the predicted number of sex selective abortions per one hundred women during their childbearing is now above seven for women with eight or more years of education.

An important conclusion of this study is that women have preferences over both the number of children they have *and* the sex composition of these children. This leads to a trade-off between the cost—both monetary and psychological—of sex selection and the cost of children. Previous research was unable to explain why higher education increased sex selection in India because it failed to tie the use of sex selection to the fertility decision. The importance of understanding this trade-off is also evident within spells. A significant feature of the new approach is that it allows for changes in the decision to use sex selection within a spell. This is most clearly seen for better-educated urban women whose only child is a girl, for whom the use of sex selection declines as the second spell becomes longer. The result is that the standard model predicts a substantially higher percentage of boys born than the hazard model, because the standard model cannot take into account censoring and changes within spells.[42]

These results cast doubts on the reliability of stated son preference as an explanation for the use of sex selection. Recent research suggests that son preference, measured as wanting more boys than girls, is decreasing over time and with higher education in India (Bhat and Zavier 2003; Pande and Astone 2007). Despite this, the use of sex selection has increased over time for exactly the group argued to show less of a stated son preference over time.

The findings also have implications for our understanding of the relation between the marriage market and sex selection. With sex selection mainly used by better-educated women, it is unlikely that higher dowries are responsible for the use of sex selection in India, unless dowries increase more than proportionally with education level and income. If dowries were an important consideration, one would expect sex selection to be used among poorer families with many girls, but there is little evidence of this. Recent data with information on both birth histories and average dowries within areas allow for a direct test of this relationship, making this a worthwhile area for future research. Another area of future research is the interaction between marriage market sex ratios and sex selection. Parents may care not only about the number and type of children they have but also about whether they will have grandchildren. If they do, then a boy could still be preferred over a girl, but a married girl would be preferred over an unmarried boy (Bhaskar 2011). The implication is that parents respond to changes in the expected sex ratio of their children's marriage market, even given their preference for sons. This hypothesis can be tested using measures of the observed distributions of boys and girls and the method proposed here.

In addition to the marriage market, there are potentially important interactions between sex selection and child outcomes. Specifically, it has been argued that more sex selection may benefit girls, the implication being that those born are more likely to be wanted by their parents. The new approach can be used to predict how likely it is that a woman has used pre-natal sex determination for a given birth based on her characteristics. This predicted probability can then be used in esti-

[42] The standard model also failed to fully captured the increased access to pre-natal sex determination for urban women and therefore underestimated the final sex ratio for third children in the period before the PNDT Act was passed.

mations of the determinants of child health to directly test whether increasing use of sex selection is beneficial for the girls who are born. It is unlikely, however, that there will be much of an effect, because sex selection is mainly used among well-educated women, who tend to have low mortality. The method can, however, also be extended to other outcomes, such as education, or the use of sex selection can be predicted for an geographical area and the predicted use for the area can then be used as an explanatory variable.

In conclusion, the results provide strong clues to how the use of sex selection will change in the future. Because lower fertility is responsible for the increase in sex selective abortions, it is likely that we will see further increases in the practice as more families want fewer children, either because of urbanization or because of increases in female education. With the already low fertility for well-educated urban women, a substantial future increase per woman is unlikely, but a higher proportion of women will belong to this group in the future. For rural, well-educated women, fertility is still falling. To the extent that it falls to the same level as for urban women, we should expect a corresponding increase in sex selection. We are beginning to see evidence of lower fertility for women with less education(i.e., below eight years of education) but a substantially number still have four or more children. Once that begins to drop to three—and even to two—we are likely to see a substantially higher use of sex selection in India.

References

Abrevaya, Jason, "Are There Missing Girls in the United States? Evidence from Birth Data," *American Economic Journal: Applied Economics*, 2009, *1* (2), 1–34.

Allison, Paul D, "Discrete-Time Methods for the Analysis of Event Histories," *Sociological Methodology*, 1982, *13*, 61–98.

Arnold, Fred, Minja Kim Choe, and T K Roy, "Son Preference, the Family-Building Process and Child Mortality in India," *Population Studies*, 1998, *52* (3), 301–315.

_ , Sunita Kishor, and T K Roy, "Sex-Selective Abortions in India," *Population and Development Review*, 2002, *28* (4), 759–785.

Ashraf, Nava, Erica Field, and Jean Lee, "Household Bargaining and Excess Fertility: An Experimental Study in Zambia," 2009.

Ben-Porath, Yoram and Finis Welch, "Do Sex Preferences Really Matter?," *The Quarterly Journal of Economics*, 1976, *90* (2), 285–307.

Bhaskar, V, "Sex Selection and Gender Balance," *American Economic Journal: Microeconomics*, 2011, *3* (1), 214–244.

_ **and Bishnupriya Gupta**, "India's missing girls: biology, customs, and economic development," *Oxford Review of Economic Policy*, 2007, *23* (2), 221–238.

Bhat, P N Mari, "Sex Ratio in India: Correspondence," *Lancet*, 2006, *367* (9524), 1725–1726.

_ **and A J Francis Zavier**, "Fertility Decline and Gender Bias in Northern India.," *Demography*, 2003, *40* (4), 637–657.

Clark, Shelley, "Son Preference and Sex Composition of Children: Evidence from India," *Demography*, 2000, *37* (1), 95–108.

Davies, James B and Junsen Zhang, "The Effects of Gender Control on Fertility and Children's Consumption," *Journal of Population Economics*, 1997, *10* (1), 67–85.

Ding, Weili and Y Zhang, "When a Son is Born: The Impact of Fertility Patterns on Family Finance in Rural China," 2009.

Dolton, Peter and Wilbert von der Klaauw, "Leaving Teaching in the UK: A Duration Analysis," *Economic Journal*, March 1995, *105* (429), 431–444.

Dreze, Jean and Mamta Murthi, "Fertility, Education, and Development: Evidence from India," *Population and Development Review*, 2001, *27* (1), 33–63.

Dyson, Tim, "The Preliminary Demography of the 2001 Census of India," *Population and Development Review*, 2001, *27* (2), 341–356.

Ebenstein, Avraham Y, "Estimating a Dynamic Model of Sex Selection in China," *Demography*, 2011, *48*, 783–811.

Edlund, Lena, "Son Preference, Sex Ratios, and Marriage Patterns," *Journal of Political Economy*, 1999, *107* (6), 1275–1304.

_ **, Hongbin Li, Junjian Yi, and Junsen Zhang**, "Sex Ratios and Crime: Evidence from China," *Review of Economics and Statistics*, September 2013.

Ejrnæs, Mette and Claus C Pörtner, "Birth Order and the Intrahousehold Allocation of Time and Education," *Review of Economics and Statistics*, November 2004, *86* (4), 1008–1019.

George, Sabu M, "Sex Ratio in India," *Lancet*, 2006, *367* (9524), 1725.

Gibson, Mhairi A. and Ruth Mace, "Strong Mothers Bear More Sons in Rural Ethiopia," *Proceedings of the Royal Society London B*, 2003, *270*, S108–S109.

Goodkind, Daniel, "On Substituting Sex Preference Strategies in East Asia: Does Prenatal Sex Selection Reduce Postnatal Discrimination," *Population and Development Review*, 1996, *22* (1), 111–125.

Gupta, Monica Das and P N Mari Bhat, "Fertility Decline and Increased Manifestation of Sex Bias in India," *Population Studies*, 1997, *51* (3), 307–315.

Han, Aaron and Jerry A Hausman, "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 1990, *5* (1), 1–28.

Jacobsen, R, H Moller, and A Mouritsen, "Natural Variation in the Human Sex Ratio," *Human Reproduction*, 1999, *14* (12), 3120–3125.

Jenkins, Stephen P, "Easy Estimation Methods for Discrete-Time Duration Models," *Oxford Bulletin of Economics and Statistics*, 1995, *57* (1), 129–138.

Jha, Prabhat, Rajesh Kumar, Priya Vasa, Neeraj Dhingra, Deva Thiruchelvam, and Rahim Moineddin, "Low male-to-female [sic] sex ratio of children born in India: national survey of 1.1 million households," *Lancet*, 2006, *367*, 211–218.

Klasen, Stephen and Claudia Wink, "A Turning Point in Gender Bias in Mortality? An Update on the Number of Missing Women," *Population and Development Review*, 2002, *28* (2), 285–312.

Lancaster, John, "The Desperate Bachelors," *The Washington Post*, 2002, *December 2*, A01.

Leung, Siu Fai, "Will Sex Selection Reduce Fertility?," *Journal of Population Economics*, 1994, *7* (4), 379–392.

Lin, Ming-Jen, Nancy Qian, and Jin-Tan Liu, "More Women Missing, Fewer Girls Dying: The Impact of Abortion on Sex Ratios at Birth and Excess Female Mortality in Taiwan," Technical Report 2008.

Merli, M Giovanna and Adrian E Raftery, "Are Births Underreported in Rural China? Manipulation of Statistical Records in Response to China's Population Policies," *Demography*, 2000, *37* (1), 109–126.

Murthi, Mamta, Anne Catherine Guio, and Jean Dreze, "Mortality, Fertility, and Gender Bias in India: A District-Level Analysis," *Population and Development Review*, 1995, *21* (4), 745–782.

Pande, Rohini P and Nan Marie Astone, "Explaining Son Preference in Rural India: The Independent Role of Structural Versus Individual Factors," *Population Research and Policy Review*, 2007, *26* (1), 1–29.

Park, Chai Bin and Nam Hoon Cho, "Consequences of Son Preference in a Low-Fertility Society: Imbalance of the Sex Ratio at Birth in Korea," *Population and Development Review*, 1995, *21* (1), 59–84.

Rasul, Imran, "Household Bargaining over Fertility: Theory and Evidence from Malaysia," *Journal of Development Economics*, 2008, *86* (2), 215–241.

Repetto, Robert, "Son Preference and Fertility Behavior in Developing Countries," *Studies in Family Planning*, 1972, *3* (4), 70–76.

Retherford, Robert D and T K Roy, "Factors Affecting Sex-Selective Abortion in India and 17 Major States," Technical Report, Mumbai, India 2003.

Sudha, S and S Irudaya Rajan, "Female Demographic Disadvantage in India 1981-1991: Sex Selective Abortions and Female Infanticide," *Development and Change*, 1999, *30* (3), 585–618.

Thomas, Jonathan M, "On the Interpretation of Covariate Estimates in Independent Competing-Risks Models," *Bulletin of Economic Research*, 1996, *48* (1), 27–39.

Trivers, R L and D E Willard, "Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring," *Science*, 1973, *179* (4068), 90–92.

Wei, Shang-Jin and Xiaobo Zhang, "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China," 2009.

Wells, Jonathan C K, "Natural Selection and Sex Differences in Morbidity and Mortality in Early Life," *Journal of Theoretical Biology*, 2000, *202* (1), 65–76.

Zeng, Yi, Ping Tu, Baochang Gu, Yi Xu, Bohua Li, and Yongpiing Li, "Causes and Implications of the Recent Increase in the Reported Sex Ratio at Birth in China," *Population and Development Review*, 1993, *19* (2), 283–302.

Zhou, Weijin, J orn Olsen, G L Nielsen, and S Sabroe, "Risk of Spontaneous Abortion Following Induced Abortion is only Increased with Short Interpregnancy Interval," *Journal of Obstetrics and Gynaecology*, 2000, *20* (1), 49–54.