

# 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  
cportner@seattleu.edu  
www.clausportner.com  
&  
Center for Studies in Demography and Ecology  
University of Washington

March 2016

---

\*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 the University of Washington Royalty Research Fund and the Development Research Group of the World Bank is gratefully acknowledged. 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. Partial support for this research 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.”

## **Abstract**

This paper addresses two main questions: what is the relationship between fertility and sex selection and how does birth spacing interact with the use of sex-selective abortions? I introduce a statistical method that incorporates how sex-selective abortions affect both the likelihood of a son and spacing between births. Using India's National Family and Health Surveys, I show that falling fertility intensifies use of sex selection, leading to use at lower parities, and longer spacing after a daughter is born. Women with 8 or more years of education, both in urban and rural areas, are the main users of sex-selective abortions and have the lowest fertility. Women with less education have substantially higher fertility and do not appear to use sex selection. Predicted lifetime fertility for high-education women declined more than 10% between 1985–1994, when sex selection was legal, and 1995–2006, when sex selection was illegal. Fertility is now around replacement level. Abortions per woman increased almost 20% for urban women and 50% for rural women between the two periods, suggesting that making sex selection illegal has not reversed its use. Finally, sex selection appears to be used to ensure one son rather than multiple sons.

JEL: J1, O12, I1 Keywords: India, prenatal sex determination, censoring, competing risk

# 1 Introduction

The males-to-females ratio at birth increased dramatically over the last three decades in India as access to prenatal sex determination expanded.<sup>1</sup> Our understanding of the use of sex selection is, however, constrained by lack of information; there are no official data on sex selection, and the few surveys that ask about use of sex selection show signs of serious under-reporting (Goodkind, 1996). With no direct information, researchers have relied on a simple method for establishing factors that affect sex selection: use the sex of children born as the dependent variable, and estimate the effects of variables on the probability of having a son.<sup>2</sup> Based on this simple method, we know that families with no sons are more likely to use sex selection the higher the parity; that use of sex selection increases with socioeconomic status, especially education; and that sex selection is more widespread in cities than in rural areas (Retherford and Roy, 2003; Jha et al., 2006; Abrevaya, 2009).<sup>3</sup>

I address two important questions that the prior literature has been unable to examine because of the simple method's limitations. First, what is the relationship between fertility decisions and use of sex selection? Second, how does birth spacing interact with use of sex-selective abortions? I introduce a novel method that directly incorporates that sex-selective abortions affect both the likelihood of a son being born *and* the duration between births. I use the method to argue that differences in fertility over time and between groups explain a substantial portion of the changes in the use of sex selection in India, and to show how spacing between births play an important role in our understanding of fertility and sex selection decisions.

We already know that fertility and son preference are related from the substantial literature on fertility stopping behavior before sex selection was available. This literature shows that families

---

<sup>1</sup> 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 saw significant changes in the sex ratio at birth (Yi and Tu Ping, 1993; Park and Cho, 1995).

<sup>2</sup> In the absence of any interventions, the probability of having a son is approximately 0.512, which is independent of genetic factors (Ben-Porath and Welch, 1976; Jacobsen, Moller and Mouritsen, 1999). With fetus sex random, a statistically significant variable indicates an association with use of sex-selective abortions. Examples of studies that have used this approach are Retherford and Roy (2003), Jha et al. (2006), and Abrevaya (2009).

<sup>3</sup> There is substantial disagreement on whether sex-selective abortion is used for the first birth (Retherford and Roy, 2003; Jha et al., 2006).

are more likely to stop child bearing after the birth of a son than after the birth of a daughter (see, for example, Das, 1987; Arnold, 1997; Clark, 2000).<sup>4</sup> It is also easy to see how declining fertility may increase use of sex selection. Take a family that wants one son. If the family is willing to have up to 4 children, the probability of having a son is more than 94 percent, even without sex selection, and that increases to almost 99 percent if the family is willing to have up to 6 children.<sup>5</sup> If the desire is instead for one son *and* a maximum of two children, there is a 24 percent chance that the family will have to resort to sex selection to achieve both targets. Despite these opposite targets, there is little empirical analysis of the link between fertility and sex selection (Park and Cho, 1995).<sup>6</sup>

Closely related to the overall fertility decisions are decisions on spacing between births. In the absence of sex-selective abortions, son preference often leads to a shorter duration until the next birth if the previous birth was a daughter (see, for example, Das, 1987; Rahman and DaVanzo, 1993; Pong, 1994; Haughton and Haughton, 1996; Arnold, 1997). The resulting shorter spacing is thought to be associated with worse health outcomes for the girls (Rutstein, 2005; Conde-Agudelo, Rosas-Bermúdez and Kafury-Goeta, 2006).

What has not previously been appreciated is that the introduction of sex-selective abortions substantially changed the relationship between son preference and birth spacing. The change happens because each abortion significantly increases the duration until the next birth. The additional duration can be divided into three parts starting from the time of the abortion. First, the uterus needs at least two menstrual cycles to recover before conception, because a less than three months space between abortion and conception substantially increases the likelihood of a spontaneous abortion (Zhou, Olsen, Nielsen and Sabroe, 2000). Secondly, the average expected time to conception is about 6 months. Finally, sex determination tests are generally reliable only from 3 months

---

<sup>4</sup> Filmer, Friedman and Schady (2009) analyse the relationship between the sex composition of previous children and subsequent fertility behavior using data from 64 countries.

<sup>5</sup> The probability of not having a son are 48.8 percent for one child, 23.8 percent for two children, 11.6 percent for 3 children, 5.7 percent for 4 children, 2.8 percent for 5 children, and 1.4 percent for 6 children.

<sup>6</sup> Simulations suggest that in Korea introduction of sex selection changed family size little, but did result in abortions of female fetuses equal to about 5 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).

of gestation onwards. The result is that each abortion delays the next birth by approximately a year.<sup>7</sup>

How spacing changes with the spread of sex-selective abortions is important for two reasons. First, if decisions on sex selection change with duration since last birth, the predicted sex ratio for a parity using the simple method will be a biased estimate of the sex ratio for the parity observed when women are done with child bearing. Families may, for example, change their decision to use sex selection between births if they do not want the space between children to be too long, or if there are concerns about possible infertility as a result of too many abortions in a row. Hence, for a given parity, the sample of women for whom we do observe a birth may behave differently from the sample of women who have not yet had a birth. This is a problem because births to women with shorter spacing are more likely to be in the survey for a given parity. I address this by incorporating spacing decisions into the estimations.

Second, spacing, by itself, can no longer be used to predict son preference. We now have a situation where longer spacing is likely observed for families with the strongest son preference because they use sex-selective abortions. To further complicate matters, we may still observe short spacing after the birth of daughter as a representation of son preference for families who either have less access to sex-selective abortions or are willing to have more children. Spacing is still important and useful in understanding sex selection and son preference, but only if combined with the outcome of that spacing—in other words the likelihood of observing a boy or a girl.

The main contribution of this paper is the introduction and application of a method that incorporates the decisions about fertility, birth spacing, and sex-selective abortions. The method allows us to better understand the relationship between these three decisions. The empirical model is a

---

<sup>7</sup> There are three well-developed technologies for determining a fetus' sex: Chorionic villus sampling (CVS), amniocentesis, and ultrasound. CVS can be applied after the shortest period of gestation (8 to 12 weeks). It is the most complicated, but also the most reliable, and an abortion can be done in the first trimester. 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. Ultrasound has the advantages of being non-invasive and relatively cheap but is associated with higher risk of faulty sex determination if done early. Generally, 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 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 6 months per abortion. For a discussion of time to conception see (Wang, Chen, Wang, Chen, Guang and French, 2003)

competing risk, non-proportional hazard model with two exit states: either a boy or a girl is born. This has three major advantages over the simple method. First and foremost, it models fertility and spacing decisions jointly with the birth outcome. Secondly, by explicitly incorporating censoring of birth spacing, it addresses any potential bias from ignoring how sex selection decisions may change with duration from the previous birth. Finally, the method allows us to predict final fertility, number of abortions, and sex ratio over a woman's reproductive life cycle. None of these is possible using the simple method.

I apply the method to birth histories from India's National Family and Health Surveys for Hindu women covering the period 1972 to 2006. There are four main findings. First, declining fertility is an important factor in the increased use of sex-selective abortions. On the one hand, women with fewer than 8 years of education have relatively high fertility and do not appear to use sex selection. On the other hand, better-educated women in both urban and rural areas have much lower fertility and are the main users of sex-selective abortions. For women with 8 or more years of education and only daughters the use is substantial: around 60 percent of the children born are boys. Furthermore, as better-educated women's fertility has declined over time, sex selection occurs for earlier and earlier parities and has correspondingly increased in intensity. There is, however, still no evidence of sex selection being used on the first birth.

Secondly, sex selection appears to be used for securing one son, rather than a large number of sons. There is only limited use of sex-selective abortions for better-educated women with one or more sons. The exception is rural women with one son and one daughter, presumably to compensate for the higher child mortality in rural areas. These results are in line with the differential stopping behavior observed in many studies before sex selection became available (Repetto, 1972; Arnold, Choe and Roy, 1998; Dreze and Murthi, 2001).

Third, the legal steps taken to combat sex-selective abortions have not been able to reverse its use. Abortion has been legal in India since 1971 and still is. In 1994, the Central Government passed the prenatal Diagnostic Techniques (PNDT) Act making determining and communicating

the sex of a fetus illegal, although the number of convictions has been low.<sup>8</sup> Although it is possible that the PDNT Act has slowed the increase in use of sex selection, the predicted number of abortions has continued to increase after the law was enacted.

Finally, the method makes it possible to estimate how many births women in the sample will have, what the final sex ratio of their children will be, and how many sex-selective abortions they will go through. Completed fertility for high-education women is predicted to now be around 2 in urban areas and 2.6 in rural areas. The number of sex-selective abortions 100 women are predicted to have during their childbearing years has risen to 9.4 for urban women and 10.5 for rural women. In other words, close to one in ten well-educated women will have at least one abortion motivated by sex selection. This is a substantial increase in sex-selective abortions compared to the period before the practice was made illegal nationwide in India in 1994. The end result is a sex ratio of 53.7% boys for well-educated urban women and 53.4% for well-educated rural women.

## 2 Theory

To illustrate how fertility, sex selection, and birth spacing interact and to inform the empirical strategy, 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 prenatal sex determination given their budget constraint, the characteristics of the children they already have, and the probability of having a son,  $\pi$ .<sup>9</sup> If parents decide to use prenatal 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 prenatal sex determination, and so on.

Parents derive utility from the number of boys,  $b$ , the number of girls,  $g$ , and parental consumption,  $c$ , realized at the end of the  $T$  periods. The utility function is separable between parental

---

<sup>8</sup> The Act is described in detail at <http://pndt.gov.in/>. It took until January 2008 for the first state, Haryana, to reach 5 convictions. Hence, private clinics operate with little risk of legal action (Sudha and Rajan, 1999). Maharashtra was the first state to pass a similar law in 1988.

<sup>9</sup> 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.

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 incorporates son preference is

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

This utility function covers a variety of different preferences from equal utility of boys and girls ( $\alpha = 0.5$ ) to an extreme son preference where there is no utility of having a daughter ( $\alpha = 1$ ).<sup>10</sup>

Parents' lifetime income is  $Y$ . Parents incur a cost,  $k$ , 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  $n = b + g$ . Prenatal sex determination carries a cost  $p_s$ , and the number of pregnancies where sex determination has taken place is  $s$ . The budget constraint is

$$c + nk + sp_s \leq Y. \quad (2)$$

This optimization problem is complex. In the first period, parents can either have no children, have a child without using prenatal sex determination, or use prenatal sex determination. If parents choose prenatal sex determination, they enter the second period with either a boy or no children. The decision then repeats itself in the second period (except for those who decided not to have a child in the first period). After only two pregnancies there are now 9 possible outcomes rather than the three if prenatal sex determination was not available.<sup>11</sup>

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

---

<sup>10</sup> Other ways of defining son preference are possible, such as wanting at least one son and not caring about sex composition after that. The basic results in this section remain the same as long as there is different utility of sons and daughters for at least some sex compositions.

<sup>11</sup> Figure A.1 shows a decision tree for two periods that illustrates this complexity.



(no pregnancy, pregnancy without prenatal sex determination, and pregnancy with prenatal 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 used for the simulations 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 prenatal sex determination ( $p \in 1, 3, 5, 30$ ), where the last value is equivalent to no prenatal 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 comes from a binomial distribution with the probability above. Appendix Table A.1 shows the simulation results for average fertility, the standard deviation of fertility, and the ratio of boys to births for selected values of son preference ( $\alpha$ ), cost of prenatal sex determination ( $p_s$ ) and cost of children ( $k$ ).

In the absence of prenatal sex determination, stronger son preference is often thought to increase overall fertility. This is based on confusing total fertility and parity progression decisions based on sex composition. For a given parity level, parents with a relatively large number of girls are, indeed, 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 in the simulation. 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 continue for much longer. The simulations show that the variance in fertility increases with stronger son preference, except when children are very expensive.

Introducing prenatal sex determination has two opposing effects on fertility in the model. On

one hand, it takes fewer births to reach a desired number of sons and this reduces fertility. On the other hand, parents no longer have to pay for unwanted children and this “wealth effect” increases fertility. The simulations show that reducing the cost of prenatal sex determination tends to reduce fertility (and lower its variance).<sup>12</sup>

For the empirical model there are two important results. First, as sex selection becomes less expensive the proportion of boys among births increases and, correspondingly, the spells between births become longer. Furthermore, the higher the parity the longer the duration between births because of the more intensive use of sex selection.

Second, 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 prenatal sex determination and possibly not having a child, the optimal decision is to reverse the decision to use sex selection. Other factors, such as disutility of too long spacing between births—although not captured by this model—have the same effect. Disutility of long spacing might arise simply from wanting children close in age, but could also come from decreasing economies of scale when children are further apart in age or from larger opportunity costs of leaving the labor market at older ages. Hence, households may trade off between their preference for a son and their preference for shorter birth spacing. Finally, concerns about possible infertility may also increase with the number of abortions, leading parents to reverse the decision to use prenatal sex determination within a spell.

### **3 Estimation Strategy**

The theoretical model has three implications for fertility decisions with access to sex-selective abortions: a higher probability that next child is a son, longer average time to next birth, and that the use of prenatal sex determination may change between one observed birth and the next. All three must be incorporated into the estimation strategy to provide unbiased and precise estimates

---

<sup>12</sup> See also the discussions in Leung (1994) and Davies and Zhang (1997).

of the use of sex-selective abortions.

The empirical model is a discrete time non-proportional competing risk hazard model with two exit states: either a boy or a girl is born.<sup>13</sup> A woman's reproductive life is divided into spells which each covers the period between births (from marriage to first birth for the first spell). For each woman,  $i = 1, \dots, n$ , the starting point for a spell is time  $t = 1$  and the spell continues until time  $t_i$  when either a birth occurs or the spell 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  $h_{ijt}$  is

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

where  $T_i$  is a discrete random variable that captures when woman  $i$ 's birth occurs. To ease presentation the indicator for spell number is suppressed. The vectors of explanatory variable  $\mathbf{Z}_{it}$  and  $\mathbf{X}_i$  include information about various individual, household, and community characteristics discussed below.

The hazard rate is specified as

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

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

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

where  $D_m = 1$  if  $t = m$  and zero otherwise. 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

---

<sup>13</sup> 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.

simply shift the hazard rates up or down independent of spell length, results in greater efficiency provided the proportionality assumption holds. The problem is that the proportional hazard model does not allow covariates to have different effects 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 prenatal 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 exacerbated by the introduction of prenatal sex determination for two reasons. First, one of the main points of this paper is that use of sex-selective abortions affects birth spacing, and use of sex selection differs across groups. Second, as discussed above, the use of sex selection may vary within a spell depending on the length of the spell.

I therefore use a non-proportional model where the main explanatory variables and the interactions between them are interacted with the baseline hazards. This is captured by the  $Z$  set of explanatory variables

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

where  $D_j(t)$  is the piece-wise linear baseline hazard and  $\mathbf{Z}_1$  captures sex composition of previous children, if any, and  $Z_2$  captures area of residence. This allows the effects of the main explanatory variables on the probabilities of having a boy, a girl, or no birth to vary over time within a spell. The use of a non-proportional specification, together with a flexible baseline hazard, also mitigates any potential effects of unobserved heterogeneity (Dolton and von der Klaauw, 1995).

The remaining explanatory variables,  $\mathbf{X}$ , enter proportionally. To further minimize any potential bias from assuming proportionality, estimations are done separately for different levels of mothers' education and for different time periods. The exact specifications and the individual variables are described below.

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.<sup>14</sup> 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 daughter 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 a variable's value increases the probability of the associated event because the probability of another event may increase even more (Thomas, 1996).

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

$$P(b_t | \mathbf{X}_k, \mathbf{Z}_{kt}, t) = \frac{\exp(D_j(t) + \alpha'_{1t} \mathbf{Z}_{kt} + \beta'_1 \mathbf{X}_k)}{1 + \sum_{l=1}^2 \exp(D_j(t) + \alpha'_{lt} \mathbf{Z}_{kt} + \beta'_l \mathbf{X}_k)}, \quad (7)$$

and the predicted probability of having a girl is

$$P(g_t | \mathbf{X}_k, \mathbf{Z}_{kt}, t) = \frac{\exp(D_j(t) + \alpha'_{2t} \mathbf{Z}_{kt} + \beta'_2 \mathbf{X}_k)}{1 + \sum_{l=2}^2 \exp(D_j(t) + \alpha'_{lt} \mathbf{Z}_{kt} + \beta'_l \mathbf{X}_k)}. \quad (8)$$

With these two probabilities it is easy to calculate, for each  $t$ , the estimated percentage of children born that are boys,  $\hat{Y}_t$ ,

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

and the associated confidence interval for given values of explanatory variables.<sup>15</sup>

<sup>14</sup> 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 4. 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.

<sup>15</sup> Within each period  $1 - P(b_t) - P(g_t)$  is the probability of not having a birth in period  $t$ .

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 estimated percent boys born by length of birth spacing using graphs.<sup>16</sup> For each graph, the extent to which the percent boys is statistically significantly above the natural sex ratio indicates the use of sex selection.

Another important aspect of the hazard model is the survival curve, which shows the probability of not having had a birth yet by spell duration. The survival curves are presented for two reasons. First, they show progression to the next birth and how quickly it occurs, if at all. Secondly, they provide “weighting” for the associated 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 a period with a flat survival curve should be given little weight because the percentage boys is based on few births. The survival curve for  $t$  is

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

or equivalently

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

For comparison the online Appendix presents the estimation strategy for the simple model. The Appendix also shows how prediction results compare between the simple model and the model suggested here.

## 4 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.<sup>17</sup> The International Institute for Population Sciences in Mumbai collected the data, which have nationwide coverage. The surveys

<sup>16</sup> The parameter estimates are available on request.

<sup>17</sup> A delay in the survey for Tripura means that NFHS-2 has a small number of observation collected in 2000.

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.

I exclude visitors to the household, as well as women married more than once, divorced, or not living with their husband, women with inconsistent information on age of marriage, and those with missing information on education. Women interviewed in NFHS-3 who were never married or where *gauna* had not been performed were also dropped. The same goes for 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 9 months. Women who reported less than 9 months between marriage and first birth remain in the sample unless they are dropped for another reason.<sup>18</sup>

Finally, I restrict the sample to Hindus, who constitute about 80 percent of India's population. If use of sex selection differ between Hindu and other religions, such as Sikhs, assuming that the baseline hazard is the same would lead to bias. The other groups are, however, each so small relative to Hindus that it is not possible to estimate different baseline hazards for each group. Furthermore, the groups are so different in terms of background and son preference that combining them into one group would not make sense.

There are four advantages to using the NFHS. First, surveys enumerators pay careful attention to spacing between births and probe for “missed” births. Second, no other surveys cover as long a period in the same amount of detail. The three NFHS rounds allow me to show the development from before sex-selective abortions were available until 2006. Third, NFHS has birth histories for a large number of women.<sup>19</sup> Finally, even if probing for missing births may not completely eliminate recall error, the overlap in cohorts covered and the large sample size make it possible to

---

<sup>18</sup> Women who report less than 9 months between marriage and first birth are retained because between 10 and 20 percent fall into this category. Although it is possible that some of these births are premature the high number of women who report a birth less than 6 months after marriage indicates that conception likely occurred before marriage in most cases.

<sup>19</sup> The Special Fertility and Mortality Survey appears to cover a much large number of households than the three rounds of the NFHS combined, but Jha et al. (2006) only use the births that took place in 1997 making their sample sizes by parity smaller than here. Their sample consists of 133,738 births of which 38,177 were first born, 36,241 second born, and 23,756 were third born. The differences in results for first born children are discussed in the online Appendix.

establish where recall error remains a problem.

Recall error arise mainly from child mortality, when respondents are reluctant to discuss deceased children.<sup>20</sup> Systematic recall error, where the likelihood of reporting a deceased child depends on the sex of the child, is especially problematic because it biases the sex ratios. Probing catches many missed births, but systematic recall error is still a potentially substantial problem. Three factors contribute to the problem here. First, girls have significantly higher mortality risk than boys. Second, son preference may increase the probability that boys are remembered relative to girls. Finally, in NFHS-1 and NFHS-2 enumerators probed only for a missed birth if the initial reported birth interval was four calendar years or more. But, given short durations between births, especially after the birth of a girl, that procedure is unlikely to pick up all missed children.

Observed sex ratios by cohort provide a straightforward way to determine whether recall error is a problem. Because prenatal sex determination techniques did not become widely available until the mid-1980s, a higher than natural sex ratio for cohorts born before that time must be the result of systematic recall error. As shown in the online Appendix, the observed sex ratio by parity becomes more male dominated the further back births took place. In addition, births in the same cohort tend to be more male dominated the more recent the survey (births in the cohort took place longer ago relative to the survey). Hence, there is evidence of recall error and the degree of recall error increases with length of time between survey and cohort.

Using cohort year of birth to analyze recall error and decide which observations to keep is, however, problematic because the year of birth for a given parity is affected by recall error; for example, a second born child listed as first born will be born later than the real first born child. Year of marriage should, however, be unaffected by recall error. Using year of marriage the basic recall error pattern remains with women married longer ago more likely to report that their child was a boy for a given parity. Similarly, comparing women married in the same period across surveys shows that women married longer ago are more likely to report having sons.

That recall error increases the longer ago somebody was married means that duration of mar-

---

<sup>20</sup> The online Appendix contains a more thorough discussion of recall error and how I address it.



riage is a better predictor of recall error than calendar year of marriage. Figure 1 shows the observed sex ratio for children reported as first born as a function of duration of marriage combining all three surveys.<sup>21</sup> The solid line is the sex ratio of 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 sex ratio (approximately 0.512). To ensure sufficient cell sizes the years are grouped in twos.

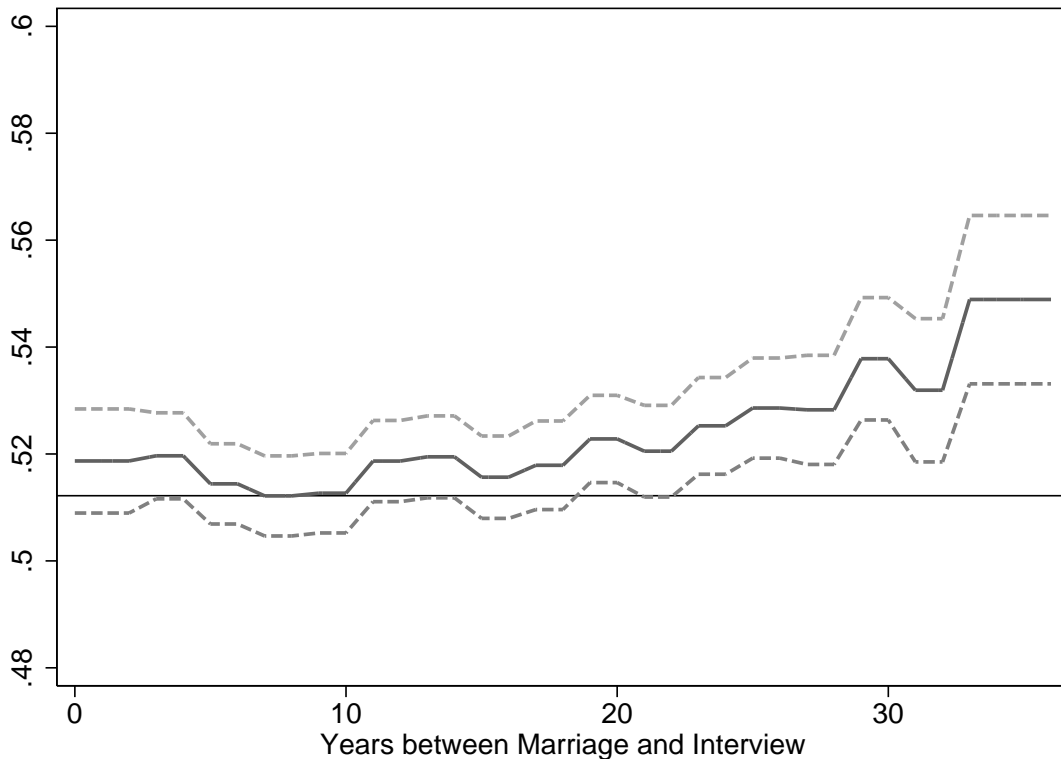


Figure 1: Ratio of boys in “first” births

Figure 1 clearly illustrates the systematic recall error problem. The observed sex ratio is increasingly above the expected value the longer ago the parents were married. The increasingly unequal sex ratio with increasing marriage duration suggests that a solution to the recall error problem is to drop women who were married “too far” from the survey year. The main problem is establishing what the best cut-off point should be. The observed sex ratio is consistently sig-

<sup>21</sup> The graph for second births shows a similar pattern. The graphs for the second births and the individual survey rounds are available upon request.

nificantly higher than the natural sex ratio from around 24 years of marriage, so one possibility is simply to drop all women married more than 24 years at the time of the survey. But, as the Appendix shows, there are differences across the three surveys and between parities. Based on the differences between surveys the cut-off points I use vary by survey round. For NFHS-1 women married 22 years or more were dropped, with the corresponding cut-off points 23 years for NFHS-2 and 26 years for NFHS-3. The final sample consists of 146,096 women, with 332,951 parity one through four births.<sup>22</sup>

#### **4.1 Spell Definition**

Spell duration is measured in quarters of a year, that is 3 month periods, hereafter referred to as quarters. The first spell begins at the month of marriage rather than 9 months after because many women report giving birth less than 9 months after they were married. For women who began living with their husbands at too young an age to conceive the starting point should ideally be first ovulation, when she becomes “at risk” for a pregnancy, rather than month of marriage. Unfortunately, information on age of menarche is only available in NFHS-1. Instead, for women who began living with their husband before age 12, I set the the month they turned 12 years of age as starting point for the first spell.

The second and subsequent spells begin 9 months after the previous birth because that is the earliest we should expect to observe a new birth. A few women report births that occurred less than 9 months after the previous birth; those women are dropped.

All spells continue until either a child is born or the spell is censored. Censoring can happen for three reasons: the survey takes place; the woman is sterilized; or the number of births observed becomes too sparse for the method to work. The timing of censoring because of too few births vary slightly by spell but is generally 21 quarters after the beginning of the spell.

I group spells into three time periods based on spell start date: 1972–1984, 1985–1994, and

---

<sup>22</sup> The online Appendix presents the results for a more restrictive definition and for the women dropped because of recall error concerns.

1995–2006. The first period covers the time before sex-selective abortions became widely available. Abortion was legalized in 1971 and amniocentesis was introduced in India in 1975, but the first newspaper reports on the availability of prenatal sex determination were not until 1982–83 (Sudha and Rajan, 1999; Bhat, 2006; Grover and Vijayvergiya, 2006). The number of clinics quickly increased, and knowledge about sex selection became widespread after a senior government 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 abortions until the prenatal Diagnostic Techniques (PNDT) act was passed in 1994. The final period is from the PNDT act until the last available survey. The PNDT act made it 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.

By dividing spells into these periods I can examine how fertility decisions and the use of sex selection have changed across the three different regimes. Note that the periods are based on the spells’ beginning year, and some spells will therefore cover two periods. A couple may, for example, be married in 1984, but not have their first child until 1986. That couple’s first spell will be in the 1972–1984 period, even though most of the spell actually falls in the 1985–1994 period. Some children born from spells that began in the 1972–1984 period may therefore have been conceived when prenatal sex determination techniques were available, which could result in evidence of sex-selective abortions even for this period. Similarly, a spell that began in the 1985–1994 period may have been partly or mostly under the PDNT act. The overall effect is to bias downwards any differences between the periods.

## **4.2 Explanatory Variables**

The explanatory variables are divided into two groups. The first group consists of variables expected to affect the shape of the hazard function: mother’s education, sex composition of previous children, and area of residence. Increasing the number of variables interacted with the baseline hazard lowers the risk of bias but requires more data to precisely estimate. I chose these variables

because the prior literature shows that they affect fertility and spacing choices and because the prior literature on sex selection indicates that these are correlated with sex selection. The second group of variables are those expected to have a proportional effect on the hazard: age of the mother at the beginning of the spell, 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.

Increasing education of mothers is strongly associated with lower fertility, with the negative effect of higher opportunity cost on fertility more than outweighing the positive effect of higher income (Schultz, 1997). Higher education should therefore be associated with higher use of sex selection. I divide women into three education groups: No education, 1 to 7 years of education, and 8 and more years of education. The models are estimated separately for each education level.

23

As discussed, the sex composition of previous children affects both the timing of fertility and the use of sex-selective abortions. Here I capture sex composition of previous children 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.<sup>24</sup> Area of residence captures both access to prenatal sex determination techniques and the cost of children. The cost of children is higher in urban areas than in rural and access to prenatal sex determination is easier; both are expected to lead to greater use of sex-selective abortions in urban than in rural areas. Because of concerns about selective migration I use where the household was living at the end of each spell.<sup>25</sup>

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

---

<sup>23</sup>Fathers' education has two opposite predicted effects: the associated higher income should increase fertility and therefore lower the pressure to use sex selection, but the higher income also makes the use of sex selection cheaper. In practice, fathers' education had little effect on the hazards and the use of sex-selective abortions and is not included.

<sup>24</sup> 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.

<sup>25</sup> As online Appendix Table D.1 shows there is little difference across education levels in migration patterns within each period and patterns across periods are also relatively stable.

be different depending on where a woman lives and the sex 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 included (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 approach 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 are 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, a 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 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.

### **4.3 Descriptive Statistics**

Appendix Table A.2 presents descriptive statistics for the spells by education level and when the spell began. 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 the survey. Hence, although about 13,000 women began the second spell there are only about 7,000 births to these women. Censoring becomes even more important for the third and fourth spells, where around 70 percent of the observations are censored. Generally, censoring increases with parity and time period. This reflects a combination of factors: timing of the surveys relative to the periods of interest, later beginning of childbearing, falling fertility,

and the longer spells from sex-selective abortions. The high number of censored observations underscore the importance of controlling for censoring when examining the relationship between fertility and sex selection.

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 groups and all three time periods. As an example, among the highly educated group for the 1995–2006 period, 51.3 percent of the children born were boys.<sup>26</sup> 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 8 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 60 percent boys in the last period, i.e. after the PNMT 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. 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. The population is also substantially better educated. Women with no education constituted almost 60 percent in the first period, but less than thirty percent in the last period. Correspondingly, in the first period just over twenty percent had 8 or more years of education, but in the last period it was almost half. Part of the increase in education is correlated with the increase in urbanization, but the proportion of better-educated women has increased substantially in the rural areas as well. Among the high education group almost 70 percent lived in urban areas during the first period but this had fallen to less than 60 percent in the last period.

The increase 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. The average number of children

---

<sup>26</sup> 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.

born to women by the time they turn 35 illustrate how strong the decline in fertility has been.<sup>27</sup> Women born in the early 1940s had on average close to 5 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 areas of residence.

## 5 Results

One of the main questions here is how fertility affects sex selection use. I start by showing the relationship between self-reported desired fertility and sex selection, which also illustrates how to interpret the results. Because of the desired fertility measure's potential shortcomings, I next examine the relationship between actual fertility and use of sex selection for better-educated women. There are two steps in this process. First, I show how fertility has declined for better-educated women with one son—a group unlikely to use sex selection, which means that sex selection will not affect the progression rates to next birth. Second, I show how, corresponding to the fertility decline, there is an increase in sex-selective abortions for better-educated women without a son. I next show how women with lower cost of children appear to follow a high fertility, low use of sex selection strategy compared to the low fertility, high use of sex selection of the better-educated women. The results all show clear evidence of son preference in fertility and sex selection use, but exactly what form son preference takes is important, so I next discuss what we can learn about son preference from the sex selection behavior. Finally, I predict lifetime fertility, number of abortions, and final sex ratios for the better-educated women.

Except for the predictions, results are presented using graphs of estimated 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 3. The “representative woman” characteristics are based on means of continuous explanatory variables and the majority category for categorical explanatory variables. Common to all “representative women” is that they do not own land and be-

---

<sup>27</sup> Figure A.2 shows the full results.

long to neither a scheduled caste nor a scheduled tribe. The graphs also show the expected natural rate of boys, approximately 51.2 percent, for comparison. Graphs for all groups and spells, even if not discussed here, are available in the online Appendix.

The survival curves show the estimated probability of not having had a birth yet by length of the spell. The survival curves are included for two reasons. First, they show the progression to the next birth by quarter and how quickly it occurs, if at all. Second, they provide “weighting” for the associated percentage boys born. The steeper the survival curve, the more weight a given spell period has because it is based on more births, whereas a flat survival curve means little weight because the percentage boys is based on few births.

## **5.1 Desired Fertility and Sex Selection**

The most direct way to understand the relationship between fertility and sex selection is to examine how use of sex-selective abortions vary with self-reported desired fertility. If lower fertility drives sex selection, women with lower desired fertility should be more likely to use sex selection than women with higher desired fertility, holding other characteristics constant. Figure 2 presents results by desired fertility for the second spell of women with 8 or more years of education whose first child was a girl.<sup>28</sup>

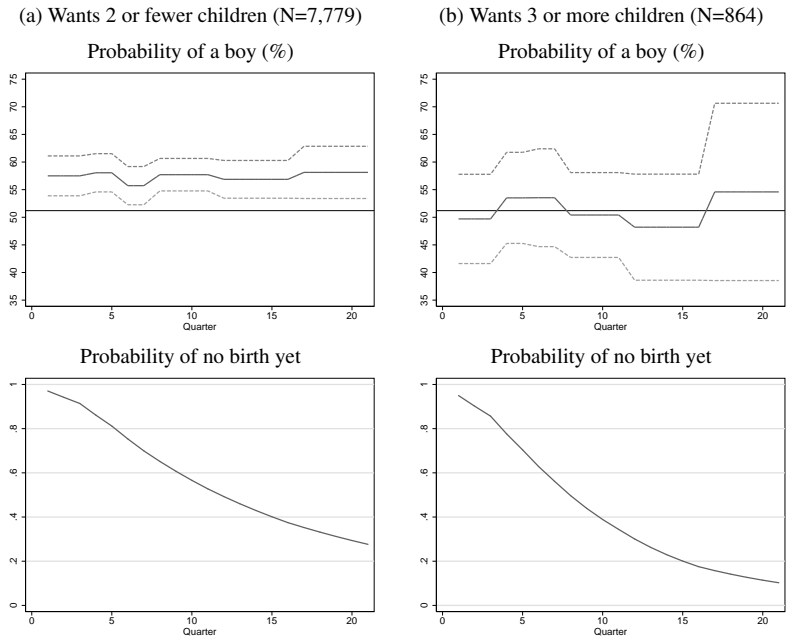
A majority of better-educated women report that they want two or fewer children, and the proportion who wants two or fewer is increasing over time. To achieve a large enough sample of women who want three or more children, I use all second spells that began in the period 1985 to 2006. This allows me to estimate non-proportional hazards by area of residence, sex of previous child, and desired fertility. For both rural and urban women, the left panel shows the results for women who want two or fewer children and the right panel shows the results for women who want three or more children.

---

<sup>28</sup> For women with living children the question asked was “If you could go back to the time you did not have any children and could choose exactly the number of children to have in your whole life, how many would that be?”, while for women with no living children the question was “If you could choose exactly the number of children to have in your whole life, how many would that be?”



## Urban



## Rural

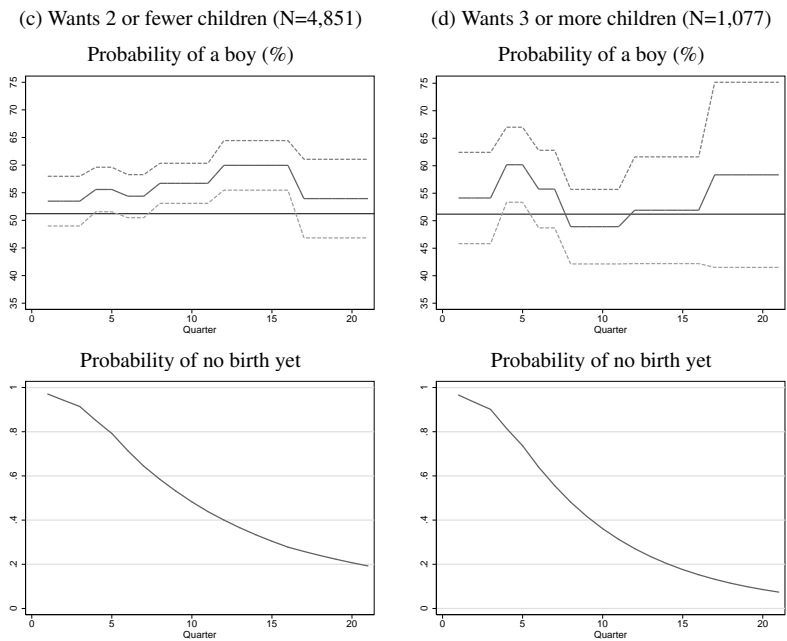


Figure 2: Desired fertility and predicted probabilities of having a boy and no birth yet from 9 months after first birth for women with 8 or more years of education by quarter (3 month period). Predictions based on first child a girl and woman age 22 at first birth. N indicates the number of women in the relevant group in the underlying samples.

Three main results stand out. First, the sex ratios by quarter for women who want 2 or fewer children show substantial evidence of sex selection, while there is little evidence among those who want 3 or more. This supports that fertility is an important factor in the use of sex selection. Second, for women who want 2 or fewer children, use of sex selection is higher among urban than rural women, but not significantly so. The proportion of boys born is between 55 and 60 percent and statistically significant higher than the natural sex ratio for all quarters for urban women, whereas for rural women the proportion is generally 2 to 3 percentage points below the urban level.<sup>29</sup>

Third, the effects of sex-selective abortions on spacing show clearly in the differences across the survival curves for the different groups. A normal spacing pattern without sex-selective abortions—where most women have the second child relatively early—leads to a survival curve with a steep slope first, followed by a leveling off. This pattern is followed by rural women who want 3 or more children. As discussed above, use of sex-selective abortions should result in substantially longer spacing between births, which makes the survival curves appear more as straight lines. Urban women who want 2 or fewer children show this pattern.

Another way to see the effects of sex-selective abortions on spacing is to compare the proportion of women who have not had a second birth yet at a specific time. Ten quarters into the spell—equivalent to 13 quarters (3¼ years) after the first birth—just below 35% of rural women who want 3 or more children have not had their second birth yet, compared with more than 45% among rural women who want 2 or fewer children. Similarly, among urban women about 35% of those who want 3 or more children have not had a second birth yet at 10 quarters after their first birth, compared with almost 55% of those who want 2 or fewer.

There are two methodological concerns with using self-reported desired fertility: it is potentially endogenous to the number of children a woman already has and may be measured with error

---

<sup>29</sup> For comparison, if 55 percent of children born in a given quarter were boys, approximately 14 percent of the female fetuses were aborted. Assume 105 boys per 100 girls born, the expected natural sex ratio. With  $b$  boys, we should expect  $b \frac{100}{105}$  girls born. If we observe  $g$  girls the number of abortions is therefore  $b \frac{100}{105} - g$  and the percent aborted female fetuses is  $\frac{b \frac{100}{105} - g}{b \frac{100}{105}} \times 100$ . With 55 percent boys we get  $\frac{55 \frac{100}{105} - 45}{55 \frac{100}{105}} \times 100 = 14.09$ . The corresponding numbers for 60 percent and 65 percent boys are approximately 30 percent and 43 percent of the female fetuses aborted.

(Jensen, 1985; Rasul, 2008; Ashraf, Field and Lee, 2009). Furthermore, from a practical standpoint there is too little variation in desired fertility to precisely estimate changes in sex selection over time. I therefore next examine whether decreases in observed fertility over time are associated with increased use of sex-selective abortions.

## 5.2 Changes in Fertility over Time

The first step when examining how declining fertility affects the use of sex selection is to show how fertility has decreased over time. The challenge is that women who use sex selection have longer birth spacing, which makes it look like they have lower fertility than they really do, and consequently would make the relationship between fertility and sex selection appear stronger than it really is. As I show below, and as suggested by the previous literature, women with one or more sons do not appear to use sex selection. Hence, these women's fertility decisions will not be affected by the availability of prenatal sex determination techniques. I therefore show fertility progression only for women who have already had at least one son, continuing to focus on the better-educated women.

Figure 3 shows the predicted progression to next birth for urban and rural women with 8 or more years of education, conditional on having at least one boy, by the starting period of the spell, and within each period by spell.<sup>30</sup> In all cases the proportion of women without a birth by the end of the spell period has increased substantial. In other words, fertility has declined over time for both urban and rural women. Although it is possible that some women will have a child after the end of each estimated spell, it is unlikely to be a substantial number given the shape of the survival curves.

---

<sup>30</sup> Could the prior births for each of these groups have been the result of sex selection? As I show in the online Appendix there is little evidence of sex selection on first births, so it is unlikely that women with a son as their first child ended up with a boy through anything other than random chance. Similarly, for the third spell group, women with two boys are unlikely to have used sex selection since there is no evidence on sex selection on first birth and no evidence of sex selection if a son is already present. It is possible that third spell women with one son and one daughter could have used sex selection in a prior spell, if they had the girl as the first child, as described below. There is, however, little difference in fertility progression for the third spell between women with two sons and women with one son and one daughter as their first children.

## Urban

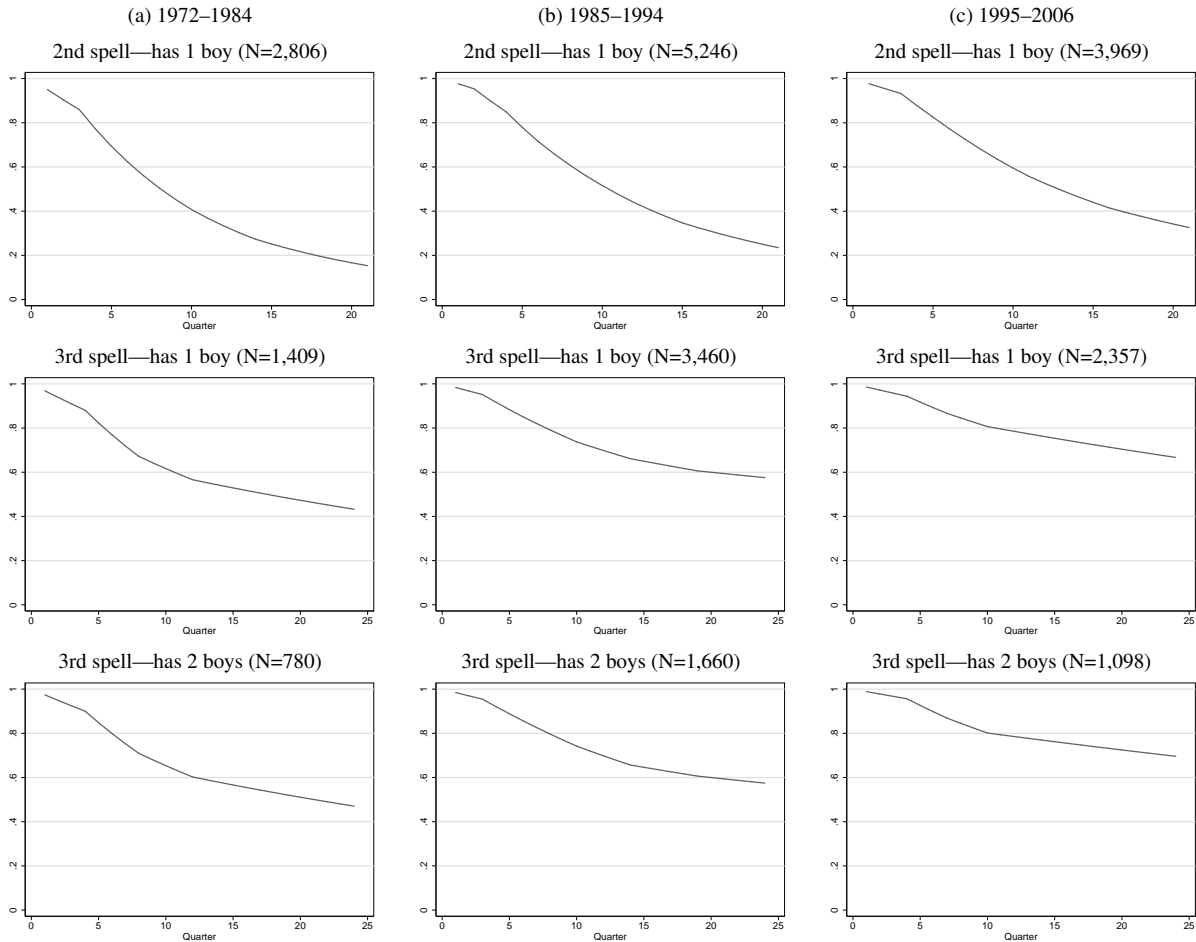


Figure 3: Predicted progression to birth for women with 8 or more years of education by quarter (3 month period) conditional on having at least one boy. Left column shows results prior to sex selection available, middle column before sex selection illegal, and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

For urban women with a son as the first child, the proportion who did not have a second child 6 years after the birth of their first child increased from around 15% in 1972–1984 period to more than 30% in the 1995–2006 period.<sup>31</sup> Similarly, the proportion who did not have a third child by 6¾ years after their second birth went from 40% in the 1972–1984 period to almost 70% in the 1995–2006 period if the woman already had one son, and from 45% to more than 70% if she had

<sup>31</sup> Recall that for the second and higher spell, the spells are estimated starting from the fourth quarter after the birth of their previous child. Hence, 3 quarters has to be added to the spell lengths in the figure to get the actual spacing.

## Rural

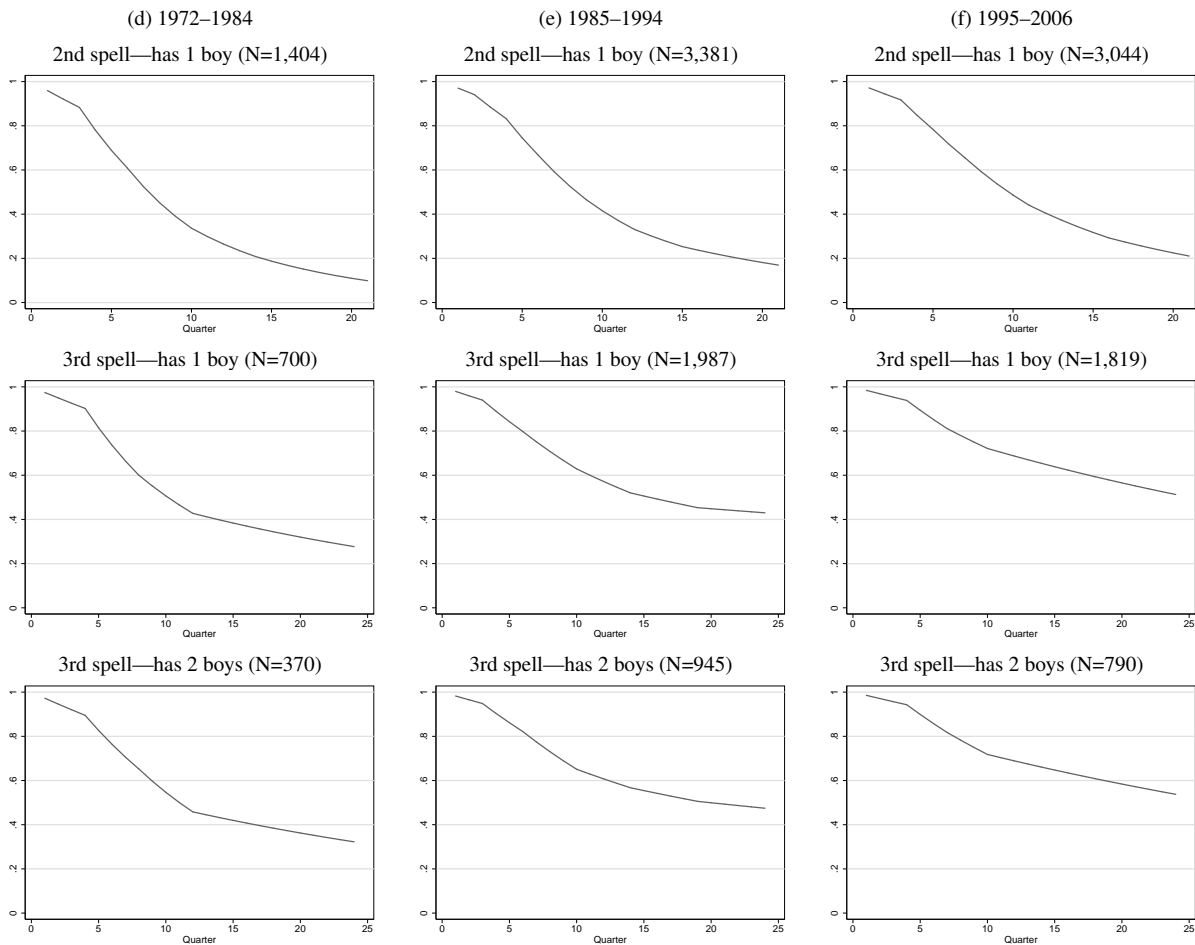


Figure 3: (Continued) Predicted progression to birth for women with 8 or more years of education by quarter (3 month period) conditional on having at least one boy. Left column shows results prior to sex selection available, middle column before sex selection illegal, and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

two sons. Rural fertility is higher, but the underlying pattern is the same. In all cases, more women finish the spell without another birth in the last period than in the first.<sup>32</sup>

Changes over time for this group of women provides an indication of whether increased availability of contraceptives affects the spacing between births over time. Increased use of modern, reversible contraceptives allows greater control over spacing, potentially making the survival curves

<sup>32</sup> The results for the fourth spell—available in the on-line Appendix—show similar results with women with at least one son experiencing large reductions in progression rates over time.

straighter if there are uniformly distributed preferences for how long spacing between births should ideally be. There are two reasons for why this is not a substantial concern here. First, use rates of modern, reversible contraceptive have not changed much. For India as a whole current use of any modern method has gone from 36.5% in NFHS-1 to 48.5% in NFHS-3, but most of this increase has come from female sterilization, which has increased from 27.4% to 37.3% over the same period (International Institute for Population Sciences (IIPS) and Macro International, 2007, p. 127). Second, survival curves in Figure 3 appear to have retained their shapes across periods to the extent possible given fertility declines, suggesting that the availability of contraceptives has not had a strong effect on the spacing between births. Since the estimations are done separately by periods the similarity of shapes come from the data, not from assumptions about functional form.

The decline in fertility over time implicitly provide information on the relative cost of having a boy versus a girl.<sup>33</sup> As discussed in Section 2, there are two opposite effects of introducing sex selection on fertility: parents can reach their desired number of sons with fewer births, which reduces fertility, but also do not have to pay for unwanted children, and this wealth effect will tend to increase fertility. The size of the wealth effect depends on the relative cost of boys and girls and the cost of sex selection; if girls are very costly and abortion cheap the wealth effect will be larger, making it more likely that fertility increases with introduction of sex selection. The cost of sex selection does, indeed, appear to be relatively low (see, for example Khanna, 1997). Hence, the decline in fertility after the introduction of sex determination is an indicator that girls are not substantially more expensive than boys.

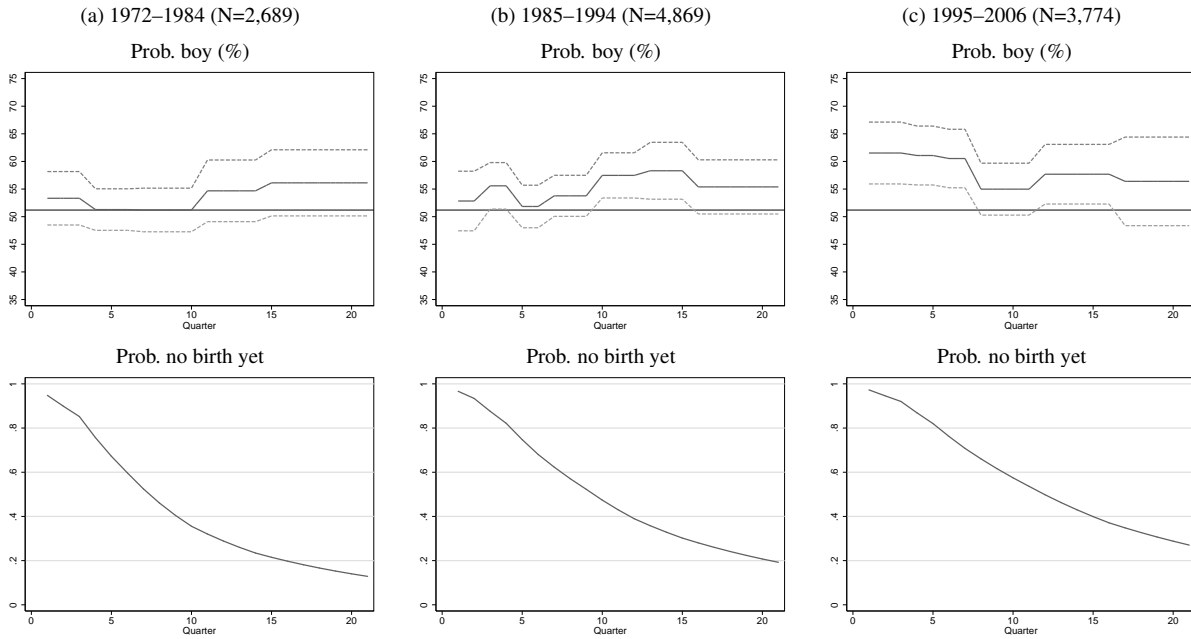
### **5.3 Effect of Declining Fertility on Sex-selective Abortions**

The second step is to show that with the fertility decline there is a corresponding increase in sex-selective abortions. If the decline in fertility drives the increase in use of sex-selective abortions, we should see an increase in sex ratios for a given spell number *and* use of sex selection for earlier spells over time. As discussed, the group most likely to use sex selection is women without sons.

---

<sup>33</sup> I am grateful to an anonymous reviewer for suggesting this point.

## Second spell for urban women—first child a girl



## Third spell—first two children girls

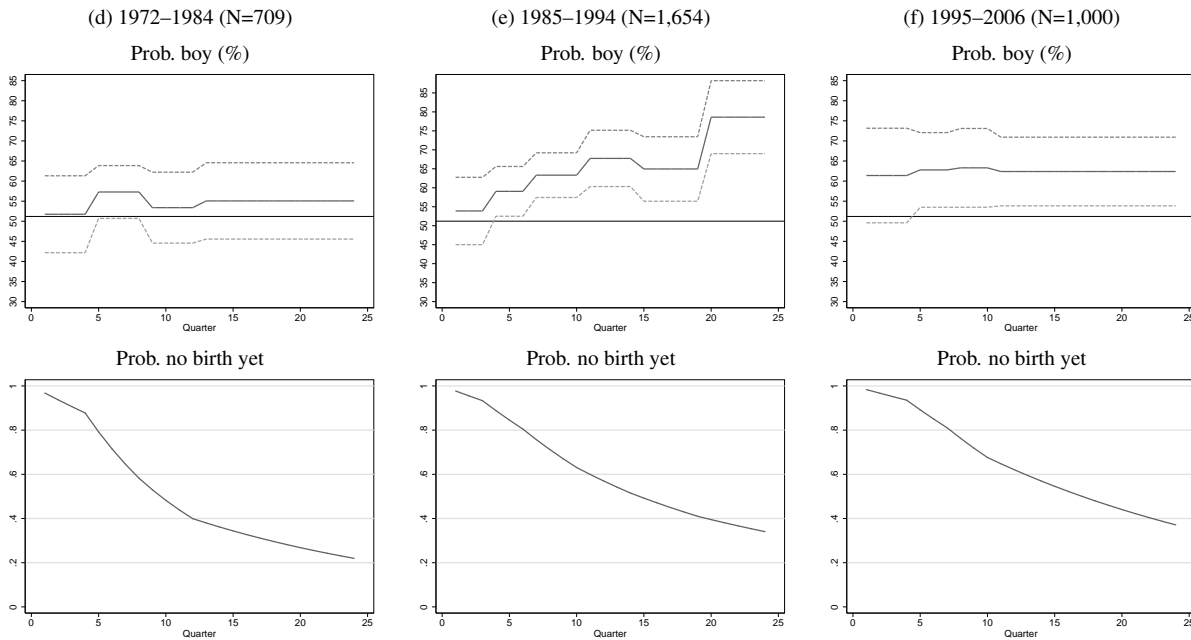
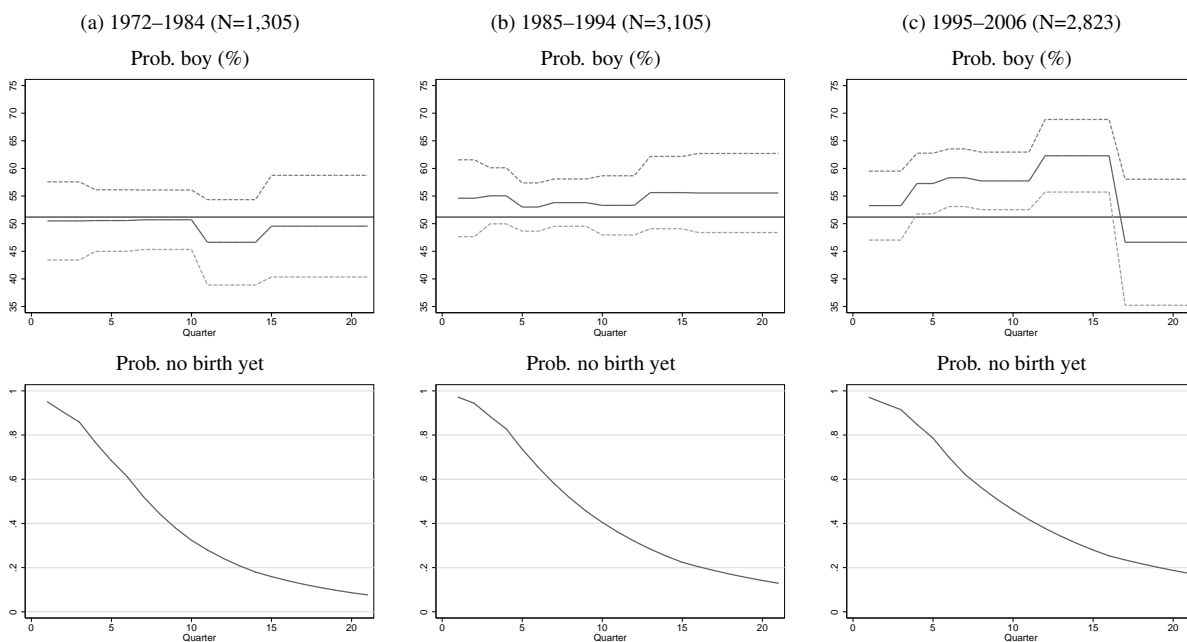


Figure 4: Predicted probability of having a boy and probability of no birth yet from 9 months after first birth for urban women with 8 or more years of education by quarter (3 month period). Predictions based on age 22 at first birth for second spell and on age 24 at second birth for third spell. Left 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.

## Second spell for rural women—first child a girl



## Third spell—first two children girls

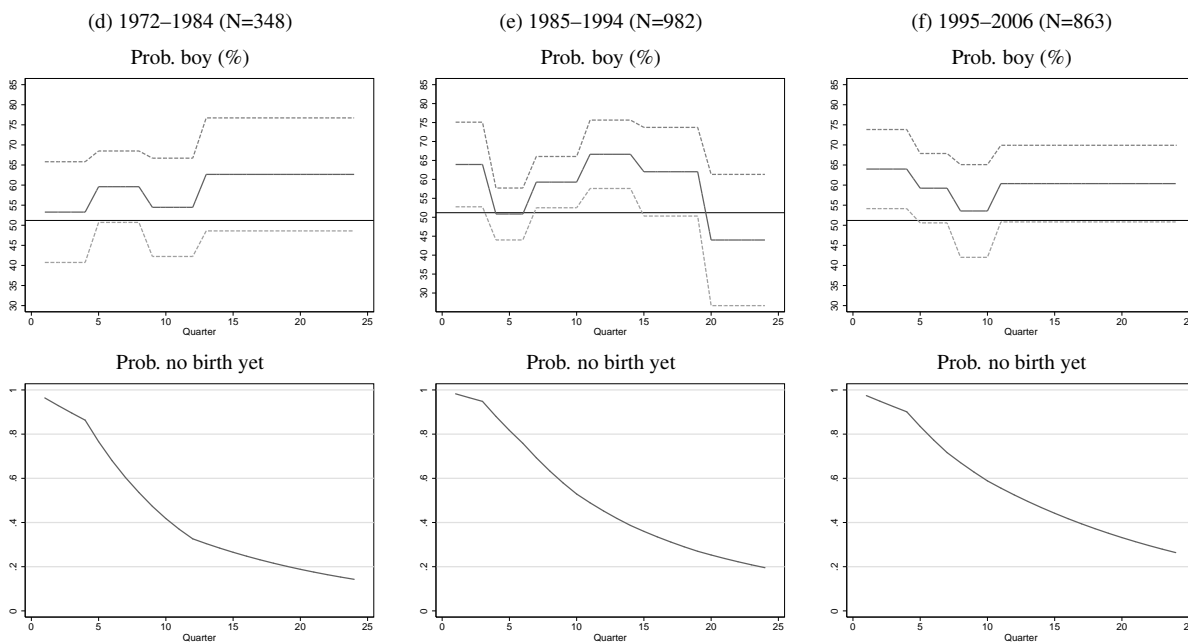


Figure 5: Predicted probability of having a boy and probability of no birth yet from 9 months after first birth for rural women with 8 or more years of education by quarter (3 month period). Predictions based on age 22 at first birth for second spell and on age 24 at second birth for third spell. Left 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.



Figures 4 and 5 therefore show the development in sex ratios by quarter for second and third spells for women without sons in urban and rural areas. Both urban and rural women follow the expected pattern, consistent with declining fertility driving the use of sex selection.

For urban women's third spell in the first period, the predicted sex ratios show no evidence of sex-selective abortions. After the introduction of pre-natal sex determination techniques there is substantial evidence of sex selection both before and after sex selection was made illegal. The increased sex ratio within the spell for the 1985–1994 period is likely the result of those women who—for one reason or another—waited longer to try for a third child having easier access to sex selection. Compared to the 1985–1994 period, there appears to be a leveling off of the sex ratio for the 1995–2006 period and the overall sex ratio may even be a little lower. Consistent with this there is little difference in the survival curves between the two periods.

Does the small difference between the two last periods for the third spell then show that banning sex selection at least stalled the increase in the use of sex selection? The results for the second spell show that this is not the case. With the falling fertility there is increased use of sex selection for the second spell. There is limited evidence of sex selection in the 1985–1994 period, which is expected given that many women were still willing to have three children. For the 1995–2006 period, however, there is a consistently high and statistically significant sex ratio. Hence, with declining fertility the use of sex selection has simply moved to an lower parity. The most likely explanation for the relative small difference between periods for the third spell is that a smaller proportion of women make it to a situation where they have two girls as their first two children in the latest period; those with the strongest preference for having a boy and for lower fertility have already used sex selection during the second spell.

Rural women follow the same general pattern as urban women, but the samples are smaller making for noisier estimates. The changes between the 1985–1994 and 1995–2006 periods for the third spell is again of special interest because it appears that the use of sex selection falls over time. But, the change for the second spell over time is even more pronounced for rural women than for urban women with very substantial increase in the use of sex selection for the second spell.

Another way to see the interaction between fertility and sex selection is to compare the survival curves across sex composition and spells. As fertility has fallen the survival curves have become straighter; the result of longer spacing arising from the use of sex selection. The proportion of women without a birth by the end of the spells have increased, but to a much smaller extent than for women with at least one son.

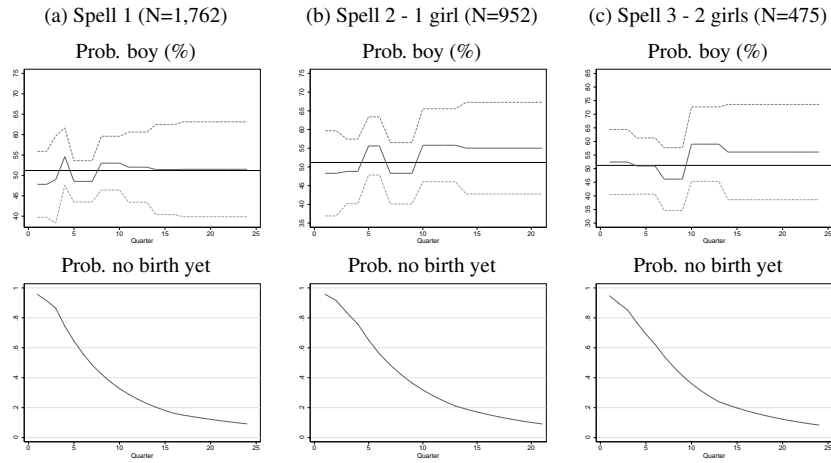
## 5.4 Fertility and Sex Selection Across Groups

As shown the decline in fertility over time is associated with both an increased use of sex-selective abortions for a given spell *and* use of sex selection for earlier spells for the women most likely to have low and falling fertility, women with 8 or more years of education. What remains is to show that women who have lower costs of children have higher fertility and consequently show less or no use of sex-selective abortions. Assuming that preferences for having one son is similar across education groups the lower cost of children makes it more attractive for less-educated women to continue having children with the aim of having at least one son instead of resorting to sex selection.

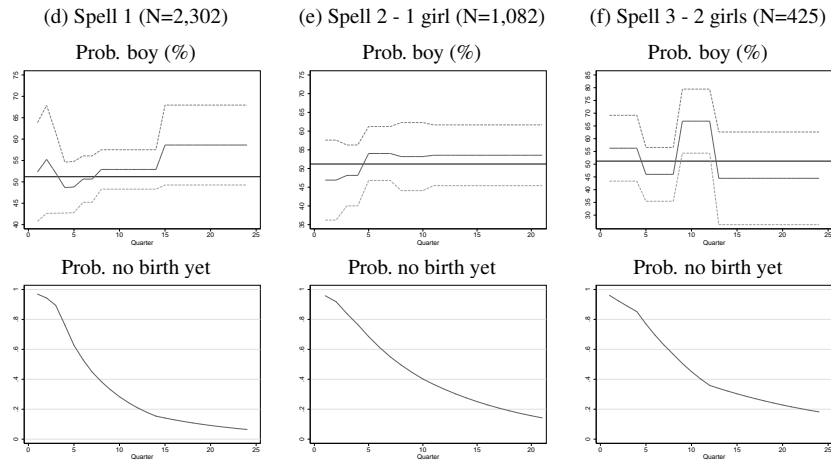
Figures 6 and 7 show sex ratios and survival curves for the three education groups for the latest period, 1995–2006, for urban and rural women. For each education group the figure shows the first spell, from marriage to first birth, the second spell for women with a girl as the first child, and, finally, the third spell for women with two girls as their first children. The second and third spells for the better-educated women are the same as in Figure 4 and 5, panels (c) and (f), and are reproduced to ease comparison.

As expected, progression rates are higher the lower the education level and higher in rural areas than in urban areas. The proportion of women with no education who end up with no birth at the end of the third spell is less than 10 % for urban women and less than 5% for rural women. For those with 1 to 7 years of education less than 20% of urban women do not go on to having a third child, while for rural women it is less than 10%. What is more, the progression rates for women with either two boys or a boy and a girl as their first children are only marginally higher than for

## No Education



## 1 to 7 Years of Education



## 8 or More Years of Education

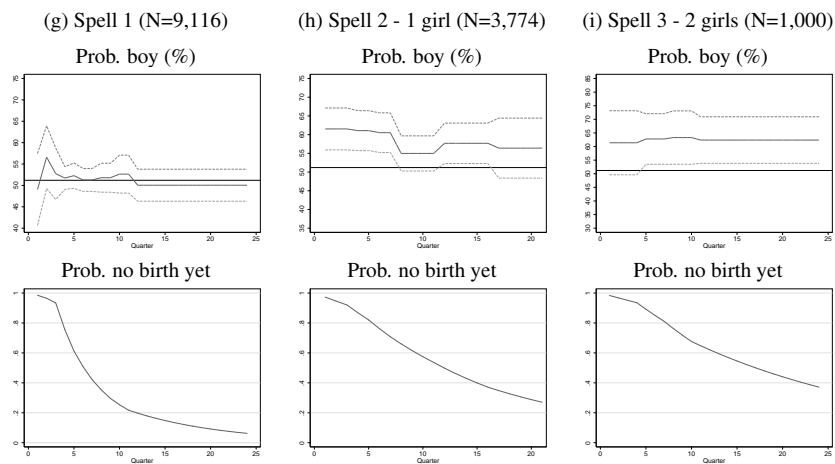
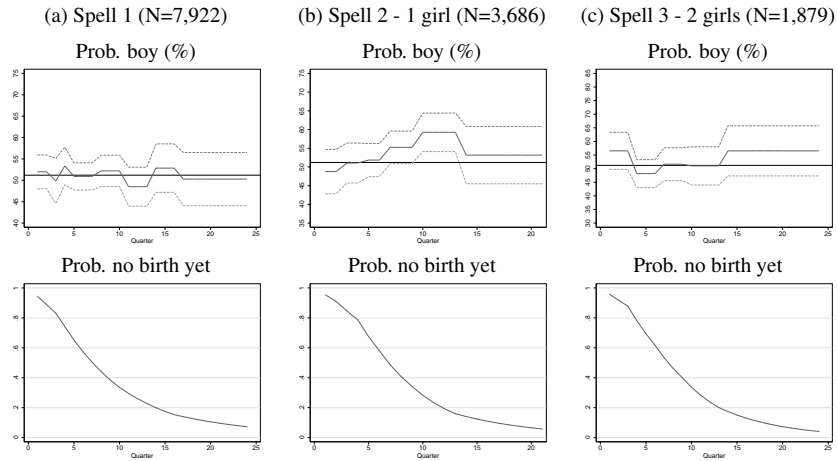
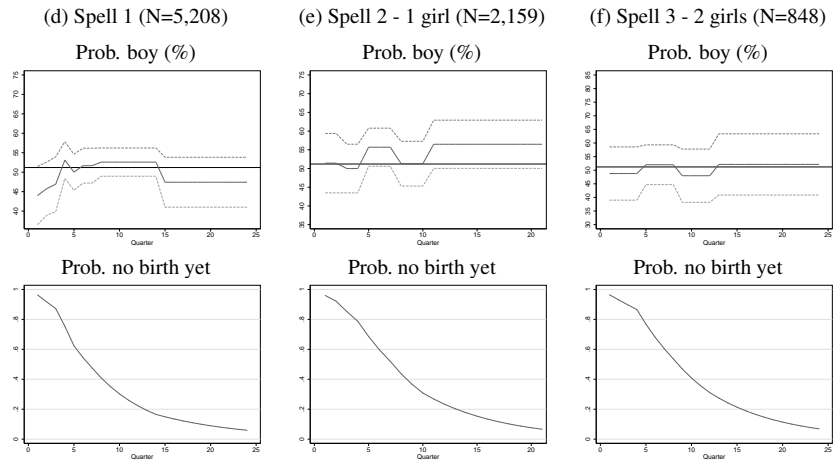


Figure 6: Predicted probability of having a boy and probability of no birth by quarter (3 month period) for urban women for 1995–2006. N is number of women in the relevant group in the underlying samples.

## No Education



## 1 to 7 Years of Education



## 8 or More Years of Education

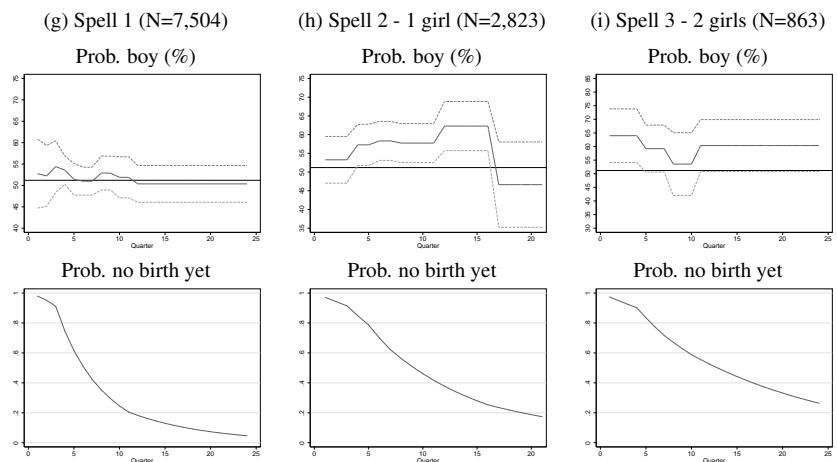


Figure 7: Predicted probability of having a boy and probability of no birth by quarter (3 month period) for rural women for 1995–2006. N is number of women in the relevant group in the underlying samples.

women with two girls.<sup>34</sup> Hence, women with less than 8 years of education still have high fertility in both urban and rural areas.

Corresponding to the higher fertility among women with less education there is little evidence of sex selection among women with less than 8 years of education. Generally the predicted sex ratios are not statistically significantly higher than the natural rate and the survival curves all show a similar pattern where most of the births occur early in the spell and there is no evidence of the changing survival curves associated with longer spells that would indicate use of sex-selective abortions.<sup>35</sup> Hence, there is evidence of two different paths to having a son: high education women have few children and substantial use of sex selection, while low education women have many children and little to no use of sex selection.

Two other important results stand out. First, there is no evidence of sex selection on first birth, no matter the education level. The sex ratios are never statistically significant different from the expected and timing of births is close to identical for all groups, with most women having their first child soon after marrying.<sup>36</sup> Second, the natural sex ratio for India appear to be in line with what is found for other countries at around 105 boys per 100 girls. For all groups for whom we would expect little to no use of sex selection and have many births to base the estimate on, such as rural women with no education and urban women with 8 or more years of education, the sex ratios almost completely coincide with the 105/100 line.

---

<sup>34</sup> In the interest of space these are shown in the online Appendix. The Appendix also shows progression rates for the fourth spell. Although samples become very small, the results correspond well. Progression rates are higher the lower the education and higher for rural women than for urban women. The samples become so small exactly for the reason discussed above: there is a less than 12% chance of not having a son by the third child.

<sup>35</sup> The exceptions are small parts of the second spell for no education women and third spell for women with 1 to 7 years of education. As the survival curves show both are based on very few births. Fourth spell results generally confirm the absence of sex selection among women with less education.

<sup>36</sup> This cast serious doubts on the data used by Jha et al. (2006) and their result that there is substantial use of sex selection on first births. As discussed above and in the online Appendix their results are likely the result of substantial recall bias.

## 5.5 One Son or Many Sons?

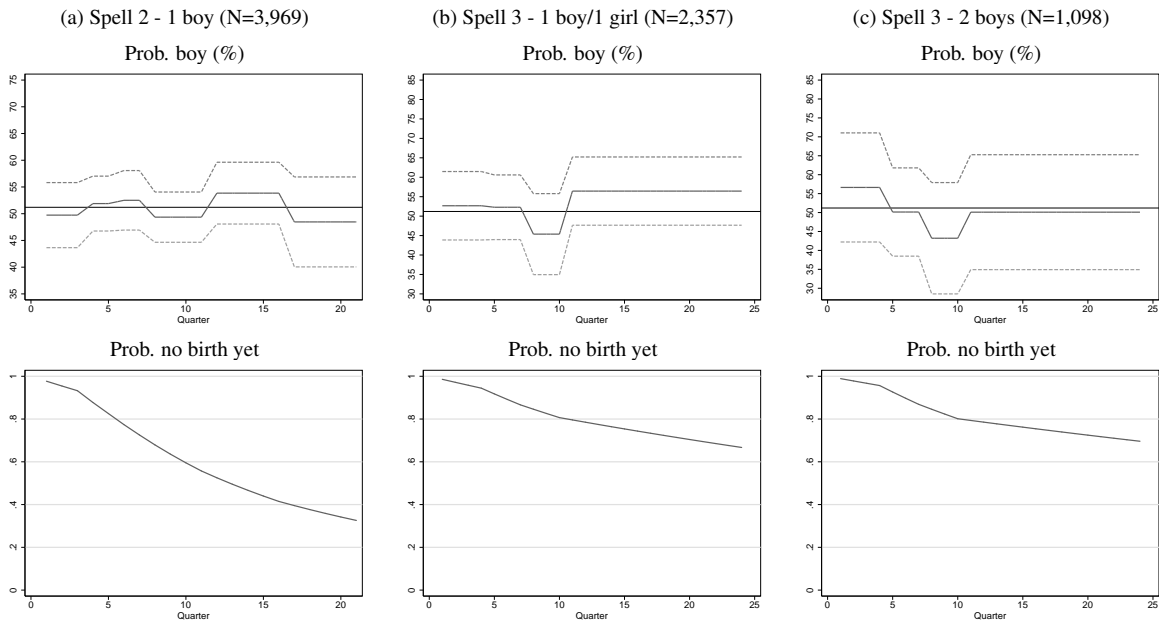
In the discussion of declining fertility I rely on sex selection not being used by families that already has one son. The extent to which this holds is of interest for two reasons. First, and foremost, it can help us understand what form son preference takes in India. This is important in itself, but especially as we try to predict what will happen to fertility and sex selection use in the future. Second, if families use sex selection to reach two or more sons progression rates for families with one son may overestimate the decline in fertility.

Figure 8 shows sex ratios and survival curves for the 1995–2006 period for urban and rural women with 8 or more years of education, conditional on having at least one son. These women are most likely to use sex selection if they have no son as shown above. Hence, if sex selection is also used to achieve more than one son, we should expect to find evidence for these women as well. The first column shows outcomes for the second spell if the first born is a son, the second column shows third spell when the first two children were a son and a daughter, and the third column shows third spell when the first two children were sons.

For the second spell the predicted sex ratios closely follow the natural sex ratio for both urban and rural women. Furthermore, the survival curves follow the expected pattern if no sex selection is taking place. There is also no evidence of sex selection for the third spell for women with two sons; the sex ratios follow the natural sex ratio for both urban and rural women. Sample sizes are, however, less than  $\frac{1}{3}$  of the sample sizes for the second spell and the progression rate to a third birth is low at only 30% for urban women and 55% for rural women, resulting in wide confidence intervals. The survival curves are particularly flat for the first year, meaning that the apparent deviations from the natural sex ratio in the beginning of the spell are based on very few births.

Even though there is no evidence of using sex selection to have an additional son when the previous children are all sons, it is possible that families want more sons than daughters and therefore use sex selection for the third spell when they already have one of each. Urban women show little evidence of this being the case. Although the sex ratio is above the natural from quarter 11 until the end of the spell, only about  $\frac{1}{3}$  of the births occur in this interval. Furthermore, the dip below

## Urban



## Rural

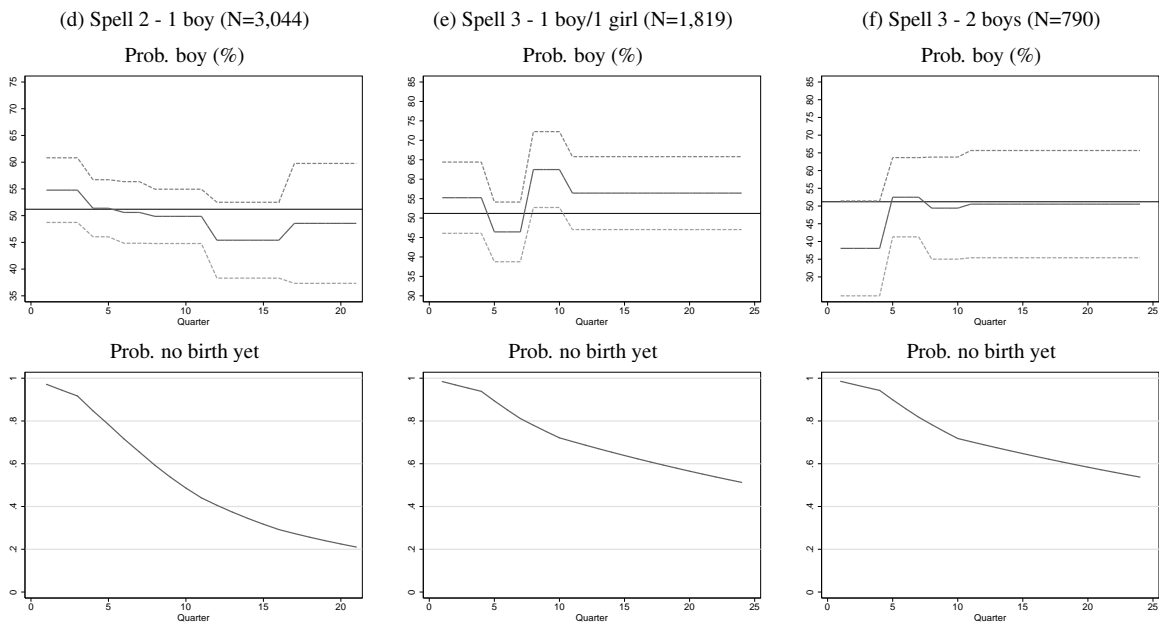


Figure 8: Predicted probability of having a boy and probability of no birth by quarter (3 month period) for women with 8 or more years of education for 1995–2006. N is number of women in the relevant group in the underlying samples.

the natural sex ratio prior to this point also accounts for approximately  $\frac{1}{3}$  of the births. Hence, these two parts cancel each other out. With the first part being to the natural sex ratio the end result is that there is little evidence for urban women using sex selection if they already have a son and a daughter.

For rural women, however, the dip and the associated number of births is not large enough to bring the overall sex ratio for the spell back to natural sex ratio. The sex ratio for the entire spell is therefore likely more male dominated than the natural rate, although it is only statistically significantly higher for two individual quarters of the spell. There are two possible explanations for rural women appearing more likely than urban women to use sex selection on the third spell if they already have a son and a daughter. First, rural women could have a stronger son preference than urban women and want two sons rather than only one. Second, mortality is higher in rural than urban areas and this higher mortality, either experienced or anticipated, make rural women use sex selection to secure a second son, essentially as an insurance. Separating these two possible explanations is difficult, but the “heir and a spare” explanation appears more likely than difference in preferences given that urban women use sex selection more than rural women for the second birth in the absence of a son. An alternative explanation is that what we are observing is simply random variation. This would explain why we get see differences in sex ratios between urban and rural women with one son across the third and fourth spells as shown in the online Appendix.<sup>37</sup>

Son preference can manifest itself through other channels than sex selection. Differential investments in education and health across boys and girls, resulting in lower education and higher mortality of girls relatively to boys, could be thought of as evidence of son preference, although Rosenzweig and Schultz (1982) argue that both may represent parents’ rational response to market opportunities rather than inherent preferences of parents. Whatever way son preference presents

---

<sup>37</sup> Urban women with a son and two daughters as their first 3 births show a sex ratio that is significantly higher than the natural, but no evidence of sex selection if they have two sons and a daughter as their first children. Rural women show only a marginally higher sex ratio with a son and two daughters and no effect at all with two sons and a daughter. These results are, however, based on very small samples. There are 590 urban women with a son and two daughters, of which less than 40% have a fourth child, 344 urban women with two sons and one daughter, and only 20% of those have a fourth child. For rural women the numbers are 564, with 40% having a fourth child, and 367 with approximately 30% having a fourth birth. This comes to approximately 300 urban births and only slightly more for rural across the two sex compositions.



itself, it appears that the dominant set of preferences as expressed through use of sex selection is for one son, rather than multiple sons.

## 5.6 Predicted Lifetime Number of Children and Abortions

One of the advantages of the method proposed here compared to the simple model is that combining spell estimates allows us to predict what fertility, number of abortions, and the sex ratios will be for women in the samples once done with child bearing. This is important for three reasons. First, we can quantify whether sex selection is intensifying or easing over time. Second, final fertility and sex ratios are what matters when examining the future impacts of sex selection in areas such as the chance of finding a spouse, savings behavior, and the potential problems associated with a surplus of bachelors (Lancaster, 2002; Ding and Zhang, 2009; Wei and Zhang, 2009; Edlund, Li, Yi and Zhang, 2013). Finally, as I show in the online Appendix and discuss above, relying only on observed sex ratios of recorded births is problematic because these are based on selection and the observed sex ratio may be biased predictors of final sex ratio when people change their use of sex selection during a spell.

Table 1 shows predicted birth rates, number of abortions, sex ratios, and completed fertility for two samples of women with 8 or more years of education: the 19,072 women who were married in 1985–1994 period and the 16,620 women who were married in the 1995–2006 period.<sup>38</sup> Predictions for each period assumes that behavior remains constant and that all spells a woman goes through are within the same period. Each woman is assumed to follow the fertility and sex selection behavior for each spell up to and including the fourth spell using the period’s estimates and her characteristics. All results are for women predicted to have one or more births.<sup>39</sup>

The predictions are done as follows. Starting with the first spell I calculate each woman’s probabilities of having a boy,  $P(b_t)$ , a girl,  $P(g_t)$ , or no birth,  $1 - P(b_t) - P(g_t)$ , in each quarter

---

<sup>38</sup> These are the women used to estimate the results for the spell from marriage to first birth above. For more information see Table A.2.

<sup>39</sup> This means that fertility predictions are larger than standard estimates, which include women with no children. Working in the opposite direction is that number of children is restricted to be at most four.

after marriage based on the period results for equations (7) and (8) using her individual values for  $\mathbf{Z}_{it}$  and  $\mathbf{X}_i$ . She then is assigned one of these outcomes—boy, girl, or no birth—based on random draws from a uniform distribution. The first quarter she has a draw that results in a birth she leaves the spell and the quarter reached is used to calculate her starting age for the second spell. If she has no draw that result in a birth by the end of the spell her childbearing stops and she does not enter the next spell. If she does have birth, the process is repeated for the second through fourth spells. Randomization is necessary because each spell’s results, except for the first, depends on the sex composition of previous children and the duration of the first spell. The results presented are averaged outcomes over all urban women and all rural women using 1,000 repetitions.

Birth/progression rates are based on the ratio of women predicted to have a birth by the end of each spell. Fertility is the average number of children born according to the randomization. With the large number of replications this number is the same as using progression rates at the end of the spell to calculate the average number of children.<sup>40</sup>

For abortions I first calculate each quarter’s probability of abortions using  $\frac{100}{105}P(b_t) - P(g_t)$ , where  $\frac{100}{105}P(b_t)$  is the probability of having a girl that should have prevailed in the absence of sex selection (0.952 girls born per boy). This probability is multiplied with the probability of not having had a birth yet at the beginning of the quarter. Abortions for each spell are then found by summing over the entire spell.<sup>41</sup> To account for random variation the abortion rate is the sum over both positive (more boys than the natural rate) and negative (more girls than the natural rate) values of the abortion calculation within the spell. Finally, the percentage of children born that are boys

---

<sup>40</sup> Given that all numbers are conditional on having at least one child, using progression rates for, say, urban women from the 1986-1994 period leads to  $1 + 1 \times 0.782 + 1 \times 0.782 \times 0.499 + 1 \times 0.782 \times 0.499 \times 0.448 = 2.3$  children.

<sup>41</sup> One way to see how this works is to imagine that we start the spell with 100 women. After the first quarter we are left with  $N_2 = (1 - P(b_1) - P(g_1)) \times 100$ , where  $N_2$  is the number of women who start the second quarter. After the second quarter we then have  $N_3 = (1 - P(b_2) - P(g_2)) \times N_2$ , etc.

Say that 65.88 women of the original 100 still have not had a birth yet by the start of quarter 6 and that the predicted probability of having a son in quarter 6 is 0.0931 and the probability of having a girl is 0.0566 (the probability of not having a birth in quarter 6 is 0.8503). For this quarter the abortion probability is then  $\frac{100}{105} \times 0.0931 - 0.0566 = 0.0321$ . In other words, the probability of having a girl should have been  $\frac{100}{105} \times 0.0931 = 0.0887$ , but only a probability of 0.0566 was predicted. Hence, in quarter 6 the predicted number of sons born is  $0.0931 \times 65.88 = 6.13$ , the predicted number of daughters born is  $0.0566 \times 65.88 = 3.73$ , and  $0.0321 \times 65.88 = 2.11$  abortions must have taken place to reach these births.

is based on the randomization of births described above.<sup>42</sup>

Table 1: Predicted Fertility Behavior, Sex Ratio, and Abortions for Women with 8 or More Years of Education

	1986-1994				1995-2006			
	Births (%) / fertility <sup>a</sup>	Abortions per 100 births <sup>b</sup>	women <sup>c</sup>	Boys (%)	Births (%) / fertility <sup>a</sup>	Abortions per 100 births <sup>b</sup>	women <sup>c</sup>	Boys (%)
<b>Urban</b>								
Spell 2	78.2	4.7	3.7	53.7	68.0	8.0	5.4	55.3
Spell 3	49.9	9.8	4.9	56.2	38.6	7.8	3.0	55.3
Spell 4	44.8	2.6	1.1	52.6	33.7	21.8	7.3	62.3
Overall <sup>d</sup>	2.3	3.4	8.0	53.1	2.0	4.6	9.4	53.7
<b>Rural</b>								
Spell 2	87.6	3.1	2.7	52.8	83.1	4.4	3.6	53.5
Spell 3	70.6	6.5	4.6	54.5	59.6	10.6	6.3	56.7
Spell 4	62.0	0.2	0.1	51.3	51.9	6.1	3.2	54.4
Overall <sup>d</sup>	2.9	2.4	6.8	52.8	2.6	4.1	10.5	53.4

**Note.** Predictions are based on estimates of equations (7) and (8). The samples consist of all women from the data with 8 or more years of education who married during the indicated period using the estimation results for that periods and the women's characteristics at the start of their marriage. All numbers presented are for women predicted to have one or more births. Birth/progression rates are based on the ratio of women predicted to have a birth by the end of each spell. Abortion rates within each spell are based on the predicted number of abortions each woman would have if she went through the entire spell. To generate conditions such as sex composition and fecundity for second and higher spells each woman is randomly assigned outcomes based on the estimation results and her characteristics as follows. For each quarter, her probabilities of having a boy, a girl, or no birth is calculated. She is assigned one of these outcomes based on a random draw from a uniform distribution. The first quarter where she has a draw that results in a birth she leaves the spell and the quarter reached is used to calculate her starting age for the subsequent spell. If she has no draw that result in a birth within a spell her childbearing stops and she does not enter the next spell. Results are averaged over 1,000 repetitions.

<sup>a</sup> For individual spells the column shows the percent of women who started the spell and ended up with a birth by the end of the spell period covered. For the "Overall" row, the number is the expected number of children born per woman over the first four spells, conditional on having at least one birth.

<sup>b</sup> For individual spells the column shows the number of abortions per 100 births predicted to take place by the end of the spell. For the "Overall" row, the number is the total number of abortion per 100 births predicted to take place over all four spells.

<sup>c</sup> For individual spells the column shows the number of abortions predicted to take place by the end of the spell per 100 women who started the spell. For the "Overall" row, the number is the total number of abortion predicted to take place over all four spells per 100 women who had at least one child.

<sup>d</sup> Covers first through fourth spells.

The results confirm that fertility has fallen over time. The percentage women with a birth by the end of each spell has decreased more than 10 percentage points between the two periods. The only exception is for rural women's second spell, where the reduction was 4.5 percentage points. The end result is that predicted completed fertility for urban women, conditional on having at least

<sup>42</sup> This number is the predicted sex ratio at birth and differ from observed sex ratios when the children are older because of mortality. On one hand, the anticipated or experienced death of a son increases the likelihood of sex selection even if parents already had a son, which would make the later observed sex ratio lower than what is predicted here. On the other hand, mortality risk is higher for girls than for boy, which would make the sex ratio more male dominated.

one child, went from 2.3 during the 1985–1994 period to just barely over 2 during the 1995–2006 period. Hence, fertility for urban women is now below replacement. Rural women showed a similar decline in completed fertility, going from 2.9 to 2.6.

Associated with the declines in progression rates and overall fertility is an increase in sex selection. For urban women the number of abortions per 100 births went from 4.7 to 8.0 for the second spell, which is a 70 percent increase. Use of sex selection has actually fallen for the third spell from 9.8 per 100 births to 7.8; this is, however, not an indication of reduced use of sex selection overall, but simply a result of fewer women making it to a third birth and those who do are now more likely to have a son because of the increased use of sex selection on the second spell. The amount of sex selection in the fourth spell confirms this: abortions per 100 births went up almost tenfold. Only 26 percent of urban women with at least one child make it to the beginning of the fourth spell, however, and less than 10 have a fourth birth, but for those who do make it to the fourth spell the pressure to ensure a son is large.

The results for rural women illustrate the two different strategies for achieving one son: either have many children or have few children and use sex selection. For the 1986–1994 period there is no use of sex selection on fourth birth, which fits with 62 percent of women end up with three or more children and a full 62 percent of women with three children have a fourth birth (recall that there is a 94 percent chance of having at least one son if you have four children). The use of sex selection on second and third spells indicates that one group of rural women follow the high fertility route, whereas the other group follow the pattern of urban women and combine restricting fertility with use of sex selection.

Overall, the number of abortions expected before childbearing is over increased substantially between the two periods. For urban women the number of abortions per 100 women have gone up by over 17 percent (8.0 to 9.4) and the number of abortions per 100 births have gone up by over 35 percent (3.4 to 4.6). For rural women, abortions per 100 women increased almost 55 percent (6.8 to 10.5), while abortions per 100 births increased over 70 percent (2.4 to 4.1).<sup>43</sup> The higher

---

<sup>43</sup> Unfortunately there are no solid official estimates to compare these numbers to. The total number of abortions per year, not just those based on son preference, vary between the official count of 600,000–700,000 and an estimate

abortion rate per woman for rural women than urban women is likely comes from a combination of higher likelihood of having an additional child and changes in use of sex selection within spell. The combined effects of the declines in fertility and increased use of sex selection are that the percentage boys increases from 53.1 to 53.7 percent for urban women, and from 52.8 to 53.4 percent for rural women.

Of particular note is the high level of abortions for the 1985–1994 period, especially for rural women. Educated rural women obviously had few constraints on access to prenatal sex determination even when the techniques were relatively new. Rural women have higher fertility 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. Furthermore, the fertility and abortion numbers for urban women for 1986-1994 and rural women for 1995-2006 are strikingly similar. Hence, as fertility continues to fall in both urban and rural areas, we are likely to see further increases in the use of sex-selective abortions.

## 6 Conclusion

If women have preferences over both the number of children they have *and* the sex composition of these children, they face a trade-off between the cost—both monetary and psychological—of sex selection and the cost of children. This paper introduces a novel approach to understanding the relationship between fertility and use of sex selection and how sex selection interact with birth spacing. The proposed method allows for joint estimation of fertility and sex-selective abortions using a non-proportional, competing risk hazard model.<sup>44</sup>

---

of 6.4 million from the 2002 Abortion Assessment Project-India. The 6.4 million abortions per year may even be an underestimate (Stillman, Frost, Singh, Moore and Kalyanwala, 2014). NFHS does ask about abortions in all three surveys, although the questions change substantially between surveys. However, the PNDT act prohibits using prenatal diagnostic tests to determine a fetus' sex and revealing a fetus' sex to parents. Hence, respondents have a strong incentive not to disclose abortions motivated by son preference. In addition, abortion are considered a taboo subject and many women are unwilling or uncomfortable reporting their abortions to government sponsored enumerators (Rossier, 2003; Stillman et al., 2014).

<sup>44</sup> This trade-off is also evident within spells. A 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.

Three results stand out. First, lower fertility is an important factor in the decision to use sex-selective abortions. Women wanting more than two births did not use sex selection on the second birth, while those who want two or fewer use sex selection intensively. In both urban and rural areas, only women with the lowest predicted fertility—those with 8 or more years of education—use sex selection, and as fertility falls they use it for earlier parities. Women with less education instead follow a high fertility strategy to ensure they have a son and do not appear to use sex selection. Previous research was unable to explain why sex selection in India only occurred among higher education women because it failed to tie the use of sex selection and the fertility decision together.

Second, 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 PNMT Act was passed. The predicted number of sex-selective abortions per 100 women during their childbearing years is now around 10 for women with 8 or more years of education in both urban and rural areas.

Third, there is little evidence that women with one or more sons use sex selection, and their probability of another birth declines substantially once they have a son. The main exception is for rural women with one son and one daughter, presumably as an insurance against mortality. Hence, parents appear to have a preference for one son rather than multiple sons.

These results imply that the way we have been measuring son preference may not be useful for understanding how people make decisions on use of sex selection. Recent research suggests that son preference in India, when measured as ideally having more boys than girls, is decreasing over time and with higher education (Bhat and Zavier, 2003; Pande and Astone, 2007).<sup>45</sup> Despite this, use of sex selection has increased and for exactly the group argued to show declining preference for sons. The culprit may be that parents want *one* son and will use sex selection intensively to get

---

<sup>45</sup> This measure of son preference is commonly used in the literature. See, for example, Clark (2000), Jensen and Oster (2009), and Hu and Schlosser (2015). In NFHS, and Demographic and Health Surveys in general, son preference is measured as part of the general fertility questions. The first question is how many children a woman would have if she could choose exactly how many. This is followed by the sex preference question: “How many of these children would you like to be boys, how many would you like to be girls and for how many would the sex not matter?”

that son, but retrain from using sex selection to have more sons, and our standard measure of son preference cannot capture that. A better approach may be to directly ask if there is a minimum number of each sex that respondents want, possibly in combination with ideal number questions. This is of more than theoretical interest. If son preference is expressed as a very strong desire for having at least one son then reductions in fertility will lead to a further increase in the use of sex selection across all groups that hold that preference.

An often-repeated, but poorly substantiated, argument for why the practice of sex selection in India is the dowry system. This explanation fits poorly with observed behavior. If true, dowries increase the cost of having girls relative to boys because you have to pay to ensure your girls are married. This implies that poorer households should be the most likely to use sex selection, and that the use of sex selection should increase the more girls they have.<sup>46</sup> Similarly, for all households there would be a financial incentive to achieve a higher number of sons than daughters. As shown here, however, sex selection is predominantly used by better educated and therefore, on average, wealthier households and only to secure one son rather than multiple sons. Finally, fertility's rapid decline after the introduction of prenatal sex determination, combined with low cost of sex selection, suggests that parents' perceived costs of boys and girls are similar.<sup>47</sup> Hence, banning dowries, as India has done, is unlikely to have little influence on the use of sex selection.

The results presented here lead to a number of important questions for future research. One is whether marriage market sex ratios affect the use of sex selection. Parents may care not only about the number and type of children they have but also whether they will have grandchildren. If they care about grandchildren, 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 a son. This hypothesis can be tested using measures of the observed distributions of boys and girls and the method proposed here.

---

<sup>46</sup> The exception would be if dowries increase more than proportionally with education level and income, but there does not appear to be evidence of this.

<sup>47</sup> See the discussion in Section 5.2 for more detail.

A second question for future research, which is especially important when designing policies aimed at reducing the use of sex selection, is who the agents are in the decisions to use sex selection. Here, I have treated the household as a single unit, but it is possible, and indeed likely, that husband and wife have different preferences over both the factors that determine use of sex selection and sex selection itself. Some of the popular press have assumed that the decision to use sex selection rests with the husband. This, however, does not fit well with the general pattern that women, on average, prefer to have fewer children than men do and that declining fertility appear to be the main driver of sex selection. A better understanding of who holds what preferences in the family would be a good first step to disentangling who makes decisions about sex selection.

A third question is whether there are interactions between sex selection and child outcomes, such a survival and education. It has been argued that access to sex selection may benefit girls because those girls who are born are more likely to be wanted by their parents. For mortality it is, however, possible that improvements are the result of a purely mechanical effect. As I have shown, increased use of sex selection lengthens the birth spacing after a girl and that by itself can improve her survival chances, even if she is no more wanted than before. The method suggested here can predict the likelihood of a woman using prenatal sex determination for a given birth based on her characteristics. This predicted probability can then be used in estimations 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 a substantial mortality effect exists since sex selection is mainly used by well-educated women, who tend to have lower mortality, but there might be effects on education investments.

In conclusion, the results provide strong clues to how sex selection use 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 already low fertility for better-educated urban women, a substantial future increase in sex-selective abortions per woman is unlikely, but a higher proportion of women will belong to this group in the future. For rural,



better-educated women fertility is still falling. To the extent that it falls to the same level as for urban women, I expect a corresponding increase in sex selection. Finally, and most importantly, we are beginning to see evidence of lower fertility for women with lower levels of education but a substantial number still have four or more children. If lower-education women hold the same preference for one son as the better-educated, once the lower-educated women's fertility begins to drop to three—and even two—we are likely to see substantially increased 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**, “Gender Preferences for Children,” Demographic and Health Surveys Comparative Studies 23, Macro International Inc, Calverton, Maryland 1997.
- , **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.
- Bhat, P N Mari**, “Sex Ratio in India,” *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.
- Conde-Agudelo, Agustin, Anyeli Rosas-Bermúdez, and Ana Cecilia Kafury-Goeta**, “Birth spacing and risk of adverse perinatal outcomes: A meta-analysis,” *JAMA*, 2006, 295 (15), 1809–1823.
- Das Gupta, Monica and P N Mari Bhat**, “Fertility Decline and Increased Manifestation of Sex Bias in India,” *Population Studies*, 1997, 51 (3), 307–315.
- Das, Narayan**, “Sex preference and fertility behavior: A study of recent Indian data,” *Demography*, 1987, 24 (4), 517–530.

- 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.
- Ebenstein, Avraham Y**, “Estimating a Dynamic Model of Sex Selection in China,” *Demography*, 2011, 48, 783–811.
- Edlund, Lena, 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.
- Filmer, Deon, Jed Friedman, and Norbert Schady**, “Development, Modernization, and Child-bearing: The Role of Family Sex Composition,” *World Bank Economic Review*, 2009, 23 (3), 371–398.
- 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.
- Grover, Anil and Rajesh Vijayvergiya**, “Sex Ratio in India,” *Lancet*, 2006, 367 (9524), 1725–1726.
- Han, Aaron and Jerry A Hausman**, “Flexible Parametric Estimation of Duration and Competing Risk Models,” *Journal of Applied Econometrics*, 1990, 5 (1), 1–28.
- Haughton, Dominique and Jonathan Haughton**, “Using a mixture model to detect son preference in Vietnam,” *Journal of Biosocial Science*, 7 1996, 28, 355–365.
- Hu, Luoia and Analía Schlosser**, “Prenatal Sex Selection and Girls’ Well-Being: Evidence from India,” *The Economic Journal*, 2015, 125 (587), 1227–1261.

- International Institute for Population Sciences (IIPS) and Macro International**, *National Family Health Survey (NFHS-3), 2005–06: India*, Vol. 1, Mumbai, India: IIPS, September 2007.
- 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.
- Jensen, Eric**, “Desired fertility, the “up to god” response, and sample selection bias,” *Demography*, 1985, *22* (3), 445–454.
- Jensen, Robert and Emily Oster**, “The Power of TV: Cable Television and Women’s Status in India,” *The Quarterly Journal of Economics*, 2009, *124* (3), 1057–1094.
- 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.
- Khanna, Sunil K**, “Traditions and Reproductive Technology in an Urbanizing North Indian Village,” *Social Science & Medicine*, 1997, *44* (2), 171–180.
- 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.
- ling Pong, Suet**, “Sex Preference and Fertility in Peninsular Malaysia,” *Studies in Family Planning*, 1994, *25* (3), pp. 137–148.
- 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.
- 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.

- Rahman, Mizanur and Julie DaVanzo**, “Gender preference and birth spacing in matlab, Bangladesh,” *Demography*, 1993, 30 (3), 315–332.
- 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.
- Rosenzweig, Mark R and T Paul Schultz**, “Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India,” *American Economic Review*, 1982, 72 (4), 803–815.
- Rossier, Clémentine**, “Estimating Induced Abortion Rates: A Review,” *Studies in Family Planning*, 2003, 34 (2), 87–102.
- Rutstein, S. O.**, “Effects of preceding birth intervals on neonatal, infant and under-five years mortality and nutritional status in developing countries: evidence from the demographic and health surveys,” *International Journal of Gynecology and Obstetrics*, Oct 2005, 89 (S1), S7–S24.
- Schultz, T Paul**, “Demand for Children in Low Income Countries,” in Mark R Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, Amsterdam: Elsevier Science B.V., 1997, pp. 349–430.
- Stillman, Melissa, Jennifer J. Frost, Susheela Singh, Ann M. Moore, and Shveta Kalyanwala**, “Abortion in India: A Literature Review,” Report, Guttmacher Institute, New York, NY December 2014.
- 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.
- Wang, Xiaobin, Changzhong Chen, Lihua Wang, Dafang Chen, Wenwei Guang, and Jonathan French**, “Conception, early pregnancy loss, and time to clinical pregnancy: a population-based prospective study,” *Fertility and Sterility*, 2003, 79 (3), 577 – 584.

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

**Yi, Zeng and Xu Yi Li Bohua Li Yongpiing Tu Ping Gu Baochang**, “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.

# A Appendix

Table A.1: Fertility and Percent Boys by Cost of Children and prenatal Sex Determination

Cost of children ( $k$ )	Cost of prenatal Sex Determination ( $p_s$ )										
	1			3			5			No Access	
	Fertility			Fertility			Fertility			Fertility	
	Mean	St. Dev.	Percent Boys <sup>a</sup>	Mean	St. Dev.	Percent Boys <sup>a</sup>	Mean	St. Dev.	Percent Boys <sup>a</sup>	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.

<sup>a</sup> Percent boys is the percent boys of all births.

Table A.2: Descriptive Statistics by Education Level and Beginning of 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	0.467 (0.499)	0.438 (0.496)	0.363 (0.481)	0.476 (0.499)	0.447 (0.497)	0.368 (0.482)	0.483 (0.500)	0.453 (0.498)	0.377 (0.485)
	Girl born	0.427 (0.495)	0.410 (0.492)	0.352 (0.478)	0.439 (0.496)	0.429 (0.495)	0.356 (0.479)	0.454 (0.498)	0.420 (0.494)	0.353 (0.478)
	Censored	0.106 (0.308)	0.152 (0.359)	0.284 (0.451)	0.085 (0.279)	0.124 (0.330)	0.276 (0.447)	0.063 (0.243)	0.127 (0.333)	0.270 (0.444)
	Urban	0.181 (0.385)	0.176 (0.381)	0.200 (0.400)	0.366 (0.482)	0.334 (0.472)	0.322 (0.467)	0.690 (0.463)	0.604 (0.489)	0.561 (0.496)
	Age	15.854 (2.482)	16.301 (2.643)	16.871 (2.858)	16.983 (2.733)	17.470 (3.041)	17.844 (3.221)	19.516 (3.331)	20.036 (3.614)	20.766 (3.916)
	Owns land	0.589 (0.492)	0.578 (0.494)	0.540 (0.498)	0.500 (0.500)	0.497 (0.500)	0.482 (0.500)	0.321 (0.467)	0.387 (0.487)	0.414 (0.493)
	Sched. caste/tribe	0.357 (0.479)	0.395 (0.489)	0.441 (0.497)	0.170 (0.376)	0.227 (0.419)	0.318 (0.466)	0.071 (0.257)	0.119 (0.324)	0.175 (0.380)
	Number of quarters	330,315	257,757	75,841	108,463	108,863	54,704	102,238	154,582	116,272
	Number of women	30,004	27,669	9,684	11,162	13,104	7,510	11,271	19,072	16,620
	Second Spell	Boy born	0.490 (0.500)	0.444 (0.497)	0.359 (0.480)	0.477 (0.500)	0.437 (0.496)	0.338 (0.473)	0.457 (0.498)	0.399 (0.490)
Girl born		0.451 (0.498)	0.413 (0.492)	0.327 (0.469)	0.462 (0.499)	0.413 (0.492)	0.319 (0.466)	0.426 (0.495)	0.351 (0.477)	0.239 (0.427)
Censored		0.059 (0.236)	0.143 (0.350)	0.314 (0.464)	0.061 (0.239)	0.150 (0.357)	0.344 (0.475)	0.117 (0.321)	0.251 (0.433)	0.476 (0.499)
One boy		0.527 (0.499)	0.515 (0.500)	0.508 (0.500)	0.519 (0.500)	0.513 (0.500)	0.505 (0.500)	0.513 (0.500)	0.520 (0.500)	0.515 (0.500)
One girl		0.473 (0.499)	0.485 (0.500)	0.492 (0.500)	0.481 (0.500)	0.487 (0.500)	0.495 (0.500)	0.487 (0.500)	0.480 (0.500)	0.485 (0.500)
Urban		0.181 (0.385)	0.180 (0.384)	0.212 (0.409)	0.370 (0.483)	0.348 (0.476)	0.345 (0.475)	0.697 (0.459)	0.628 (0.483)	0.578 (0.494)
Age		17.774 (2.743)	18.357 (3.054)	18.870 (3.284)	18.620 (2.864)	19.200 (3.206)	19.581 (3.373)	20.988 (3.371)	21.589 (3.621)	22.252 (3.901)
First spell length		25.630 (20.345)	29.145 (25.044)	29.133 (27.216)	21.924 (18.142)	24.513 (22.027)	24.787 (23.204)	20.608 (15.371)	22.154 (18.651)	22.217 (19.176)
Owns land		0.590 (0.492)	0.572 (0.495)	0.528 (0.499)	0.497 (0.500)	0.492 (0.500)	0.468 (0.499)	0.313 (0.464)	0.369 (0.483)	0.402 (0.490)
Sched. caste/tribe		0.353 (0.478)	0.390 (0.488)	0.430 (0.495)	0.164 (0.371)	0.218 (0.413)	0.304 (0.460)	0.068 (0.251)	0.110 (0.313)	0.169 (0.374)
Number of quarters	176,227	231,546	68,455	67,432	103,656	49,236	77,416	164,752	119,864	
Number of women	21,171	28,251	9,428	8,209	12,449	6,542	8,204	16,601	13,610	

Note. Means without parentheses and standard deviation in parentheses. Interactions between variables, baseline hazard dummies and squares not shown. Quarters refer to number of 3 month periods observed.



Table A.2: (Continued) Descriptive Statistics by Education Level and Beginning of 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 (0.500)	0.425 (0.494)	0.319 (0.466)	0.455 (0.498)	0.388 (0.487)	0.261 (0.439)	0.353 (0.478)	0.264 (0.441)	0.163 (0.369)	
	Girl born	0.442 (0.497)	0.393 (0.488)	0.306 (0.461)	0.423 (0.494)	0.357 (0.479)	0.239 (0.426)	0.313 (0.464)	0.214 (0.410)	0.128 (0.334)	
	Censored	0.080 (0.271)	0.182 (0.386)	0.376 (0.484)	0.121 (0.326)	0.255 (0.436)	0.500 (0.500)	0.334 (0.472)	0.521 (0.500)	0.709 (0.454)	
	Two boys	0.276 (0.447)	0.257 (0.437)	0.243 (0.429)	0.248 (0.432)	0.245 (0.430)	0.232 (0.422)	0.266 (0.442)	0.244 (0.429)	0.238 (0.426)	
	One boy, one girl	0.490 (0.500)	0.502 (0.500)	0.506 (0.500)	0.508 (0.500)	0.501 (0.500)	0.515 (0.500)	0.489 (0.500)	0.510 (0.500)	0.527 (0.499)	
	Two girls	0.234 (0.424)	0.241 (0.428)	0.252 (0.434)	0.243 (0.429)	0.254 (0.435)	0.253 (0.435)	0.245 (0.430)	0.247 (0.431)	0.235 (0.424)	
	Urban	0.179 (0.383)	0.182 (0.386)	0.208 (0.406)	0.378 (0.485)	0.354 (0.478)	0.343 (0.475)	0.693 (0.461)	0.646 (0.478)	0.568 (0.495)	
	Age	19.978 (2.895)	20.693 (3.196)	21.278 (3.551)	20.794 (2.937)	21.453 (3.239)	21.998 (3.618)	22.963 (3.405)	23.825 (3.748)	24.548 (4.067)	
	First spell length	24.547 (18.832)	27.575 (22.779)	27.099 (24.364)	20.971 (16.875)	23.901 (20.378)	23.532 (21.155)	19.991 (14.646)	21.246 (16.835)	21.236 (17.363)	
	Owns land	0.602 (0.490)	0.575 (0.494)	0.538 (0.499)	0.503 (0.500)	0.503 (0.500)	0.490 (0.500)	0.320 (0.467)	0.369 (0.483)	0.421 (0.494)	
	Sched. caste/tribe	0.344 (0.475)	0.392 (0.488)	0.432 (0.495)	0.159 (0.366)	0.218 (0.413)	0.298 (0.457)	0.067 (0.251)	0.107 (0.310)	0.169 (0.375)	
	Number of quarters	109,026	206,988	70,391	43,434	87,537	41,400	49,799	126,733	79,068	
	Number of women	13,196	25,135	9,355	4,960	9,775	5,034	4,316	10,688	7,927	
	Fourth Spell	Boy born	0.463 (0.499)	0.384 (0.486)	0.291 (0.454)	0.403 (0.491)	0.347 (0.476)	0.233 (0.423)	0.310 (0.462)	0.227 (0.419)	0.170 (0.376)
		Girl born	0.408 (0.491)	0.363 (0.481)	0.268 (0.443)	0.391 (0.488)	0.303 (0.460)	0.198 (0.399)	0.298 (0.458)	0.197 (0.397)	0.113 (0.316)
Censored		0.130 (0.336)	0.254 (0.435)	0.441 (0.497)	0.206 (0.405)	0.350 (0.477)	0.569 (0.495)	0.392 (0.488)	0.576 (0.494)	0.717 (0.450)	
Three boys		0.134 (0.341)	0.124 (0.329)	0.111 (0.315)	0.107 (0.309)	0.105 (0.307)	0.102 (0.303)	0.106 (0.308)	0.101 (0.302)	0.082 (0.275)	
Two boys, one girl		0.374 (0.484)	0.357 (0.479)	0.346 (0.476)	0.348 (0.476)	0.323 (0.468)	0.308 (0.462)	0.358 (0.480)	0.304 (0.460)	0.289 (0.453)	
One boys, two girls		0.362 (0.481)	0.389 (0.488)	0.400 (0.490)	0.400 (0.490)	0.413 (0.492)	0.423 (0.494)	0.393 (0.489)	0.439 (0.496)	0.469 (0.499)	
Three girls		0.129 (0.335)	0.130 (0.337)	0.142 (0.350)	0.145 (0.353)	0.159 (0.365)	0.167 (0.373)	0.143 (0.350)	0.156 (0.363)	0.159 (0.366)	
Urban		0.171 (0.377)	0.181 (0.385)	0.197 (0.398)	0.363 (0.481)	0.346 (0.476)	0.325 (0.469)	0.670 (0.470)	0.601 (0.490)	0.511 (0.500)	
Age		21.926 (3.019)	22.885 (3.332)	23.608 (3.711)	22.611 (2.894)	23.519 (3.364)	24.317 (3.759)	24.257 (3.171)	25.472 (3.702)	26.061 (4.143)	
First spell length		23.281 (17.481)	26.609 (21.556)	26.460 (23.272)	20.495 (15.546)	23.069 (19.078)	24.215 (21.244)	19.106 (13.991)	20.826 (16.277)	21.423 (16.585)	
Owns land		0.610 (0.488)	0.587 (0.492)	0.557 (0.497)	0.523 (0.500)	0.530 (0.499)	0.509 (0.500)	0.339 (0.474)	0.414 (0.493)	0.466 (0.499)	
Sched. caste/tribe		0.338 (0.473)	0.403 (0.491)	0.442 (0.497)	0.154 (0.361)	0.224 (0.417)	0.297 (0.457)	0.085 (0.280)	0.113 (0.316)	0.193 (0.395)	
Number of quarters		54,446	151,285	56,756	18,852	49,108	23,209	14,287	43,339	23,586	
Number of women		6,832	18,436	7,637	2,201	5,564	2,768	1,347	3,933	2,459	

Note. Means without parentheses and standard deviation in parentheses. Interactions between variables, baseline hazard dummies and squares not shown. Quarters refer to number of 3 month periods observed.

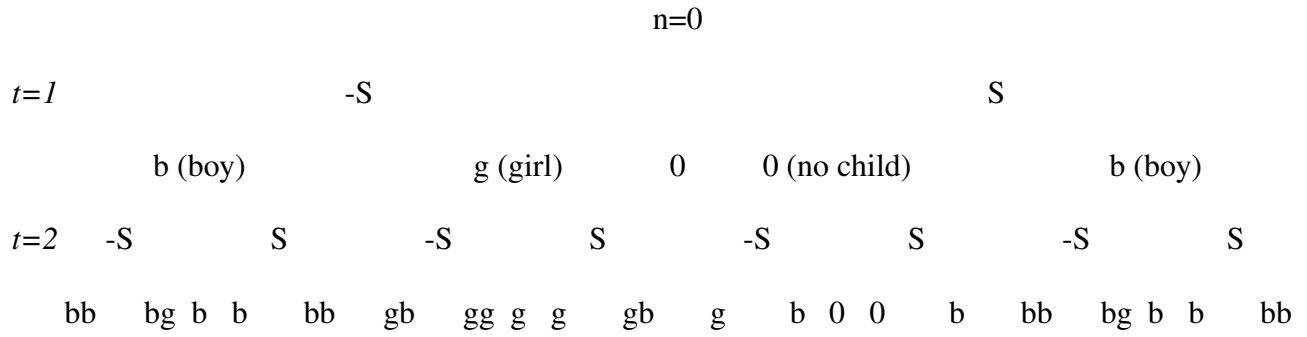


Figure A.1: Decision tree for a two period model with prenatal sex screening, where -S is no prenatal sex determination and S is sex determination

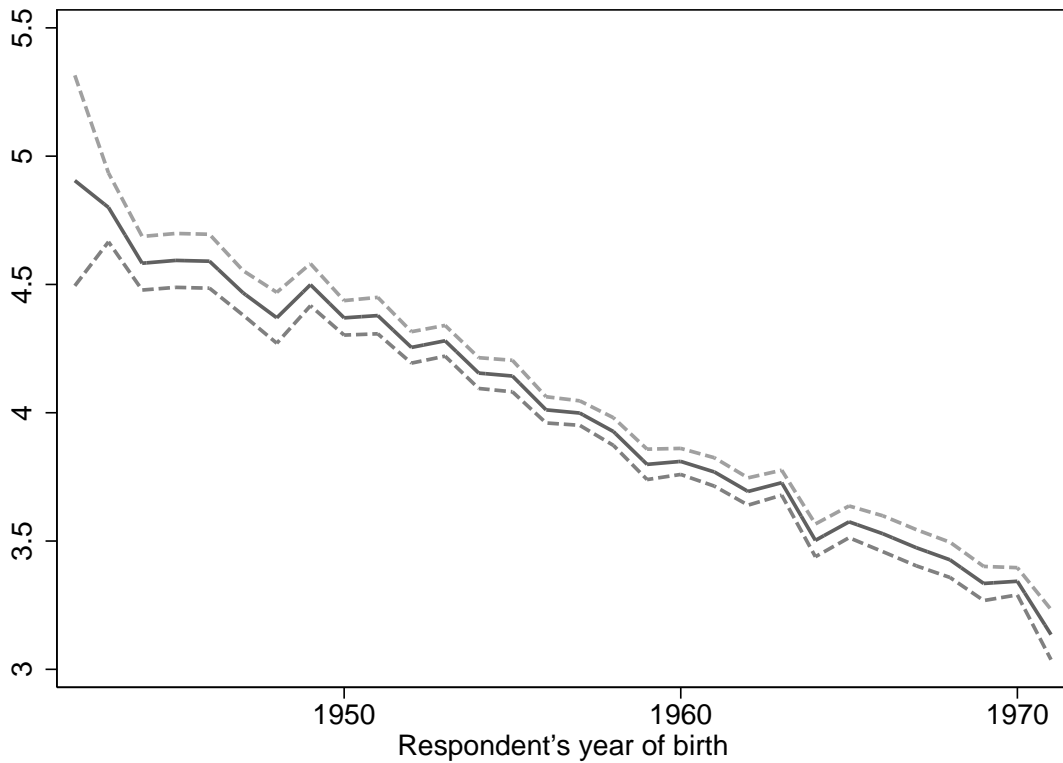


Figure A.2: Children ever born at age 35