## A Recall Error and the Sex Ratio

Since the proposed method relies both on spacing between births and birth outcomes, the quality of the results depends on the reliability of the birth histories provided by the respondents. An important concern here is child mortality. A respondent may not provide information about deceased children because of the painful nature of the topic. Furthermore, the likelihood of reporting a deceased child may depend on sex of the child. Systematic recall error, where children of a specific sex are missed non-randomly, is especially problematic because it directly biases the sex ratios results.

NFHS enumerators probe for any missed births, although the method depends on the survey. ${ }^{1}$ Probing catches many originally missed births but systematic recall error based on son preference may still be a problem. First, strong son preference lead to significantly higher mortality for girls than boys. Secondly, for children who died son preference makes it more likely that parents will remember boys than girls. Finally, in the absence of sex selective abortions, son preference lead parents to have the next birth sooner if the last child was a girl than if it was a boy. If this girl subsequently dies she is more likely to be missed because probing for missed births is only done for long intervals. ${ }^{2}$

To assess the extent of systematic recall error Table A. 1 shows sex ratios of children recorded as first-born and second-born by year of birth together with tests for whether the observed sex ratio is significantly higher than the natural sex ratio. Births are combined into five-year cohorts to achieve sufficient power for a test of differences in sex ratios. Prenatal sex determination techniques did not become widely available until the mid-1980s and any significant deviation from the natural sex ratio before that time is therefore likely the result of recall error. Furthermore, births and deaths

[^0]that took place longer ago may be more likely to be subject to recall error than more recent events. The three NFHS surveys make it possible to examine whether recall error increases with time by comparing sex ratios for births that took place during the same years but were surveyed at different times. Table A. 1 therefore also shows tests for whether more recent surveys have a higher sex ratio for the cohort than older surveys.

The "first-born" sex ratios illustrate the systematic recall error problem well. According to the birth histories in NFHS-1 almost 55 percent of children born between 1960 and 1964 were boys. Given that this is more than two decades before sex selection techniques became available in India the most likely explanation for the skewed sex ratio is that some children listed as first-borns were not, in fact, the first children born in their families. Instead, for a substantial proportion of families the first-born was a girl who died and was skipped when enumerators asked about birth history.

As expected, the difference between the observed sex ratio and the natural sex ratio is less pronounced the closer to the survey date the cohort is. The observed sex ratio for children born just prior to the NFHS-1 survey and listed as first-born is 0.517 , which is not statistically significant different from the natural sex ratio. The same general pattern holds for NFHS-2 and NFHS-3, with cohorts further away from the survey date more likely to have a sex ratio skewed male.

Second births show a pattern very similar to that for first births. An interesting difference is that there is evidence of a U-shaped relationship between time and sex ratio for second births. Cohorts furthest away from the survey year show the highest sex ratios, but declines to the natural sex ratio in the mid-1980s, and the sex ratio is then significantly higher again for more recent births. This is in line with the results in the paper that show that sex selective abortions take place on lower parity births as desired fertility declined.

Finally, across surveys the same cohort tends to show a higher sex ratio the more recent the survey (births in the cohort took place longer ago), especially for second-borns. ${ }^{3}$ Despite this, few cohorts show significantly different sex ratios across surveys, most likely because of a lack of power.

[^1]Table A.1: Observed sex ratios by year of birth in five-year cohorts

|  | Listed as first-born |  |  |  | Listed as second-born |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \hline \text { NFHS-1 } \\ \text { 1992-1993 } \end{gathered}$ | $\begin{gathered} \text { NFHS-2 } \\ \text { 1998-1999 } \end{gathered}$ | $\begin{gathered} \text { NFHS-3 } \\ \text { 2005-2006 } \end{gathered}$ | Diff. test ${ }^{\text {a }}$ | $\begin{gathered} \hline \text { NFHS-1 } \\ \text { 1992-1993 } \end{gathered}$ | $\begin{gathered} \hline \text { NFHS-2 } \\ \text { 1998-1999 } \end{gathered}$ | $\begin{gathered} \hline \text { NFHS-3 } \\ \text { 2005-2006 } \end{gathered}$ | $\begin{aligned} & \text { Diff. } \\ & \text { test } \end{aligned}$ |
| 1960-1964 | $\begin{gathered} 0.5473^{* * *} \\ (0.0004) \\ {[2,673]} \end{gathered}$ | ${ }_{[.]}^{(.)^{2}}$ | (.) <br> [.] |  | $\begin{gathered} 0.5172 \\ (0.4382) \\ {[1,189]} \end{gathered}$ | ${ }_{[.]}^{(.)}$ | ${ }_{[.)_{[.]}}$ |  |
| 1965-1969 | $\begin{gathered} 0.5274^{* *} \\ (0.0306) \\ {[5,429]} \end{gathered}$ | $\begin{gathered} 0.5538^{* * *} \\ (0.0002) \\ {[1,997]} \end{gathered}$ | (.) [.] | A | $\begin{gathered} 0.5327^{* *} \\ (0.0124) \\ {[3,888]} \end{gathered}$ | $\begin{gathered} 0.5244 \\ (0.3310) \\ {[595]} \end{gathered}$ | (.) [.] |  |
| 1970-1974 | $\begin{gathered} 0.5354^{* * *} \\ (0.0001) \\ {[7,763]} \end{gathered}$ | $\begin{gathered} 0.5341^{* * *} \\ (0.0019) \\ {[5,546]} \end{gathered}$ | $\begin{gathered} 0.5411 \\ (0.1208) \\ {[523]} \end{gathered}$ |  | $\begin{gathered} 0.5212 \\ (0.1497) \\ {[6,249]} \end{gathered}$ | $\begin{gathered} 0.5313^{* *} \\ (0.0248) \\ {[3,499]} \end{gathered}$ | $\begin{gathered} 0.6160^{* *} \\ (0.0143) \\ {[125]} \end{gathered}$ | BC |
| 1975-1979 | $\begin{gathered} 0.5207 \\ (0.1258) \\ {[8,749]} \end{gathered}$ | $\begin{gathered} 0.5151 \\ (0.4669) \\ {[7,478]} \end{gathered}$ | $\begin{gathered} 0.5255^{*} \\ (0.0920) \\ {[3,762]} \end{gathered}$ |  | $\begin{gathered} 0.5185 \\ (0.2477) \\ {[7,819]} \end{gathered}$ | $\begin{gathered} 0.5262^{* *} \\ (0.0340) \\ {[6,220]} \end{gathered}$ | $\begin{gathered} 0.5384^{* *} \\ (0.0225) \\ {[1,811]} \end{gathered}$ | B |
| 1980-1984 | $\begin{gathered} 0.5210^{*} \\ (0.0872) \\ {[11,147]} \end{gathered}$ | $\begin{gathered} 0.5244^{* *} \\ (0.0273) \\ {[9,709]} \end{gathered}$ | $\begin{gathered} 0.5280^{* * *} \\ (0.0094) \\ {[7,718]} \end{gathered}$ |  | $\begin{gathered} 0.5236^{* *} \\ (0.0361) \\ {[9,950]} \end{gathered}$ | $\begin{gathered} 0.5236^{* *} \\ (0.0492) \\ {[8,478]} \end{gathered}$ | $\begin{gathered} 0.5249^{*} \\ (0.0628) \\ {[5,532]} \end{gathered}$ |  |
| 1985-1989 | $\begin{gathered} 0.5187 \\ (0.1916) \\ {[11,276]} \end{gathered}$ | $\begin{gathered} 0.5132 \\ (0.6153) \\ {[11,048]} \end{gathered}$ | $\begin{gathered} 0.5118 \\ (0.7091) \\ {[9,422]} \end{gathered}$ |  | $\begin{gathered} 0.5157 \\ (0.4099) \\ {[10,296]} \end{gathered}$ | $\begin{gathered} 0.5194 \\ (0.1695) \\ {[9,857]} \end{gathered}$ | $\begin{gathered} 0.5165 \\ (0.3716) \\ {[8,022]} \end{gathered}$ |  |
| 1990-1994 | $\begin{gathered} 0.5174 \\ (0.3263) \\ {[6,544]} \end{gathered}$ | $\begin{gathered} 0.5185 \\ (0.2006) \\ {[11,653]} \end{gathered}$ | $\begin{gathered} 0.5183 \\ (0.2207) \\ {[10,601]} \end{gathered}$ |  | $\begin{gathered} 0.5120 \\ (0.6570) \\ {[5,869]} \end{gathered}$ | $\begin{gathered} 0.5166 \\ (0.3430) \\ {[10,653]} \end{gathered}$ | $\begin{gathered} 0.5289^{* * *} \\ (0.0030) \\ {[9,299]} \end{gathered}$ | BC |
| 1995-1999 | (.) [.] | $\begin{gathered} 0.5222^{*} \\ (0.0788) \\ {[8,642]} \end{gathered}$ | $\begin{gathered} 0.4980 \\ (0.9998) \\ {[11,111]} \end{gathered}$ |  | (.) [.] | $\begin{gathered} 0.5230^{*} \\ (0.0664) \\ {[7,990]} \end{gathered}$ | $\begin{gathered} 0.5262^{* *} \\ (0.0103) \\ {[9,956]} \end{gathered}$ |  |
| 2000-2006 | (.) [.] | (.) <br> [.] | $\begin{gathered} 0.5135 \\ (0.5978) \\ {[13,413]} \end{gathered}$ |  | (.) <br> [.] | (.) [.] | $\begin{gathered} 0.5292^{* * *} \\ (0.0008) \\ {[11,645]} \end{gathered}$ |  |

Note. Sample consists of Hindu women only. First number in cell is ratio of boys to girls. Second number in parentheses is p-value for the hypothesis that observed sex ratio is greater than 105/205 using a binomial probability test (bitest in Stata 13) with asterisks indicating significance as follows: * sign. at $10 \% ; * *$ sign. at $5 \%$; *** sign. at $1 \%$. Third number in square brackets is number of observations.
${ }^{\text {a }}$ Test (prtest in Stata 13) whether recall error increases with time passed, which would manifest itself in a higher sex ratio for a more recent survey than an earlier for the same cohort. A: Cohort sex ratio significantly larger in NFHS-2 than NFHS-1 at the 10 percent level. B: Cohort sex ratio significantly larger in NFHS-3 than NFHS-1 at the 10 percent level. C: Cohort sex ratio significantly larger in NFHS-3 than NFHS-2 at the 10 percent level.

Using year of birth to analyze recall error is problematic because year of birth for a given parity is affected by recall error; for example, a second born child listed as first born will naturally be born later than the real first born child. Recorded year of marriage should, however, be affected
neither by parental recall error nor use of sex selective abortions. Table A. 2 therefore shows sex ratios of children recorded as first-born and second-born by year of parents' marriage, together with tests for whether the observed sex ratio is significantly higher than the natural sex ratio and whether more recent surveys show a higher sex ratio for the cohort than older surveys.

The basic recall error pattern remains with women married longer ago more likely to report that their first-born is a boy. If anything the pattern is more consistent when using year of marriage than year of birth. Before the introduction of sex selection there are only two instances across the two parities where the observed sex ratio is lower for women married longer ago than for women married closer to the survey date. ${ }^{4}$ Similarly, comparing women married in the same five-year period across surveys shows that women married longer ago are more likely to report having a son.

The relationship between length of marriage and recall error can also seen in Figure A.1, which shows the observed sex ratio for children reported as first born as a function of duration of marriage at the time of the survey combining all three surveys. 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. In line with the results from Table A.2, the observed ratio of boys is increasingly above the expected value the longer ago the parents were married. ${ }^{5}$

The increasingly unequal sex ratio with increasing marriage duration suggests that a solution to the recall error problem is to drop observations for women who were married "too far" from the survey year. The main problem is establishing what the best cut-off point should be. As Table A. 2 shows there are differences in recall error across the three surveys and between parities. One possibility is simply to drop all women married in cohorts where the observed sex ratio is substantially higher than the natural sex ratio, but even this is not straightforward. For "first" born

[^2]Table A.2: Observed sex ratios by year of parents' marriage in five-year cohorts

|  | Listed as first-born |  |  |  | Listed as second-born |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{gathered} \text { NFHS-1 } \\ \text { 1992-1993 } \end{gathered}$ | $\begin{gathered} \text { NFHS-2 } \\ \text { 1998-1999 } \end{gathered}$ | $\begin{gathered} \text { NFHS-3 } \\ \text { 2005-2006 } \end{gathered}$ | Diff. test ${ }^{\text {a }}$ | $\begin{gathered} \text { NFHS-1 } \\ \text { 1992-1993 } \end{gathered}$ | $\begin{gathered} \text { NFHS-2 } \\ \text { 1998-1999 } \end{gathered}$ | $\begin{gathered} \text { NFHS-3 } \\ \text { 2005-2006 } \end{gathered}$ | Diff. test ${ }^{\text {a }}$ |
| 1960-1964 | $\begin{gathered} \hline 0.5368^{* * *} \\ (0.0002) \\ {[6,200]} \end{gathered}$ | (.) [.] | (.) [.] |  | $\begin{gathered} \hline 0.5274^{* *} \\ (0.0239) \\ {[6,001]} \end{gathered}$ | (.) [.] | (.) [.] |  |
| 1965-1969 | $\begin{gathered} 0.5333^{* * *} \\ (0.0011) \\ {[6,707]} \end{gathered}$ | $\begin{gathered} 0.5450^{* * *} \\ (0.0000) \\ {[4,268]} \end{gathered}$ | ${ }_{[.]}^{(.)}$ |  | $\begin{gathered} 0.5277^{* *} \\ (0.0181) \\ {[6,449]} \end{gathered}$ | $\begin{gathered} 0.5401^{* * *} \\ (0.0005) \\ {[4,142]} \end{gathered}$ | (.) [.] |  |
| 1970-1974 | $\begin{gathered} 0.5244^{* *} \\ (0.0386) \end{gathered}$ $[8,154]$ | $\begin{gathered} 0.5222 \\ (0.1120) \\ {[6,546]} \end{gathered}$ | $\begin{gathered} 0.5271 \\ (0.1372) \\ {[1,973]} \end{gathered}$ |  | $\begin{gathered} 0.5182 \\ (0.2651) \\ {[7,843]} \end{gathered}$ | $\begin{gathered} 0.5235^{*} \\ (0.0805) \\ {[6,308]} \end{gathered}$ | $\begin{gathered} 0.5384^{* *} \\ (0.0194) \\ {[1,913]} \end{gathered}$ | B |
| 1975-1979 | $\begin{gathered} 0.5272^{* * *} \\ (0.0062) \\ {[9,829]} \end{gathered}$ | $\begin{gathered} 0.5202 \\ (0.1477) \\ {[8,644]} \end{gathered}$ | $\begin{gathered} 0.5316^{* * *} \\ (0.0050) \\ {[5,798]} \end{gathered}$ | C | $\begin{gathered} 0.5157 \\ (0.4208) \\ {[9,295]} \end{gathered}$ | $\begin{gathered} 0.5183 \\ (0.2499) \\ {[8,294]} \end{gathered}$ | $\begin{gathered} 0.5278^{* *} \\ (0.0242) \\ {[5,608]} \end{gathered}$ | B |
| 1980-1984 | $\begin{gathered} 0.5151 \\ (0.4616) \\ {[10,814]} \end{gathered}$ | $\begin{gathered} 0.5142 \\ (0.5322) \\ {[9,924]} \end{gathered}$ | $\begin{gathered} 0.5195 \\ (0.1877) \\ {[8,316]} \end{gathered}$ |  | $\begin{gathered} 0.5198 \\ (0.1514) \\ {[9,788]} \end{gathered}$ | $\begin{gathered} 0.5171 \\ (0.3178) \\ {[9,409]} \end{gathered}$ | $\begin{gathered} 0.5187 \\ (0.2332) \\ {[7,906]} \end{gathered}$ |  |
| 1985-1989 | $\begin{gathered} 0.5176 \\ (0.2753) \end{gathered}$ | $\begin{gathered} 0.5209^{*} \\ (0.0925) \\ {[10,990]} \end{gathered}$ | $\begin{gathered} 0.5095 \\ (0.8425) \\ \\ \hline 9717 \end{gathered}$ |  | $\begin{gathered} 0.5131 \\ (0.5918) \end{gathered}$ | $\begin{aligned} & 0.5194 \\ & (0.1651) \end{aligned}$ | $\begin{gathered} 0.5267^{* *} \\ (0.0105) \\ {[9.077]} \end{gathered}$ | B |
| 1990-1994 | $\begin{gathered} 0.5180 \\ (0.3836) \end{gathered}$ | $\begin{gathered} 0.5182 \\ (0.2282) \\ {[10,619]} \end{gathered}$ | $\begin{gathered} 0.5125 \\ (0.6692) \\ {[10,574]} \end{gathered}$ |  | $\begin{array}{r} 0.4366 \\ (0.9740) \\ {[142]} \end{array}$ | $\begin{gathered} 0.5191 \\ (0.2097) \\ {[7,994]} \end{gathered}$ | $\begin{gathered} 0.5251^{* *} \\ (0.0199) \\ {[9,600]} \end{gathered}$ | AB |
| 1995-1999 | (.) <br> [.] | $0.5243^{*}$ $(0.0845)$ <br> [5,081] <br> [5,081] | $\begin{gathered} 0.5017 \\ (0.9965) \\ {[10,960]} \end{gathered}$ |  | (.) <br> [.] | $\begin{gathered} 0.5613^{* * *} \\ (0.0015) \\ {[1,019]} \end{gathered}$ | $\begin{gathered} 0.5312^{* * *} \\ (0.0009) \\ {[8,969]} \end{gathered}$ |  |
| 2000-2006 | (.) <br> [.] | (.) <br> [.] | $\begin{gathered} 0.5169 \\ (0.3280) \\ {[9,210]} \end{gathered}$ |  | (.) <br> [.] | (.) <br> [.] | $\begin{gathered} 0.5247 \\ (0.1243) \\ {[3,316]} \end{gathered}$ |  |

Note. Sample consists of Hindu women only. First number in cell is ratio of boys to girls. Second number in parentheses is p-value for the hypothesis that observed sex ratio is greater than 105/205 using a binomial probability test (bitest in Stata 13) with asterisks indicating significance as follows: * sign. at $10 \% ; * *$ sign. at $5 \%$; *** sign. at $1 \%$. Third number in square brackets is number of observations.
${ }^{\text {a }}$ Test (prtest in Stata 13) whether recall error increases with time passed, which would manifest itself in a higher sex ratio for a more recent survey than an earlier for the same cohort. A: Cohort sex ratio significantly larger in NFHS-2 than NFHS-1 at the 10 percent level. B: Cohort sex ratio significantly larger in NFHS-3 than NFHS-1 at the 10 percent level. C: Cohort sex ratio significantly larger in NFHS-3 than NFHS-2 at the 10 percent level.
this would imply dropping all women married before 1980, but for "second" born it would be imply dropping all married before 1970 for NFHS-1, before 1975 for NFHS-2, and before 1980

NFHS-3. Because of the differences between "first" and "second" born in the different surveys I


Figure A.1: Sex ratio of "first" births by duration of marriage
vary the cut-off point by survey round.
NFHS-3 is the most straightforward in that the implied cut-off point is the same for both first and second born. Based on Table A.2, women in NFHS-3 who were married 26 years or more before the survey date are dropped. For NFHS-1 and NFHS-2 the trade-off is between ensuring enough information and not bias the results unduly. Given that sex selective abortions are more likely in higher parity birth I base the cut-off point on the second births instead of the first births. Hence, for NFHS-1 women who were married 22 years or more before the survey date were dropped and for NFHS-2 women who were married 23 year or more before the survey date were dropped. The final sample consists of 146,096 women, with 332,951 parity one through four births.

## A. 1 Sensitivity of Results to Recall Error

The last three sections in this Appendix show results for the three education groups for three samples: the preferred sample used in the paper (Figures E. 1 through E.21), a more restricted sample where only women married less than 19 years are included (Figures F. 1 through F.15), and the sample of women dropped because of potential recall error bias, i.e. all women dropped from the data to reach the preferred sample (Figures G. 1 through G.15). For the "dropped" sample it is only possible to estimate sex ratios for the period before sex-selective abortions became available. The period covered is different from the other two samples because the births for the sample takes place over a much longer period; the sample covers spells that begin between 1950 and 1984.

The different results for women with no education illustrates the potential problems with recall error. Figures E.1, F.1, and G. 1 show the different results for women with no education across the three samples. The preferred and the restricted samples show almost identical results with the main difference that the restricted sample has wider confidence intervals because of the smaller sample size. The results from the dropped sample, on the other hand, show sex ratios that are statistically significantly above the natural sex ratio in many periods, which is consistent with recall error. Furthermore, the survival curves are substantially closer to being linear indicating that observed births took place further away from the time of marriage as expected if the couple had a girl first but counted a subsequent son as their first-born.

The second spell results demonstrate the trade-off inherent in dealing with recall error. While the results for women with no education are similar across the preferred and the restricted samples in rural areas (Figures E. 5 and F.5), the results for urban women with no education in the restricted sample are much less stable and have substantially wider confidence intervals than the preferred sample (Figures E. 4 and F.4). This happens because the restricted sample is down to less than 900 urban women, compared to close to 1,600 urban women in the preferred sample.

An interesting difference to this pattern is for the third spell for women with 8 or more years of education. This group shows no evidence on an unequal sex ratio for the "dropped" sample in either urban or rural areas (Figures G. 14 and G.15). However, for the preferred sample there are
consistently more boys than girls for urban women when the first two children are either two girls or one girl and one boy (Figure E.14). Furthermore, the sex ratios become even more unequal when restricting the sample to women married less than 19 years (Figure F.14), although the sample sizes also become substantially smaller. This is because spell periods are based on the starting year of the spell. Hence, some of the pregnancies in the 1972-1984 period will have been exposed to the availability of prenatal sex determination. For Figure E.14(a) 25 percent of observed births occur in 1986 or after, meaning that conception took place in the 1985-1994 period, while for Figure E.14(d) the number is 29 percent. Hence, it is possible that the unequal sex ratios observed are, indeed, evidence of sex-selective abortions.

## B Sex Ratios at First Birth

I present a separate analysis of first spell births for three reasons. First, previous research claims that the largest number of missing girls is for first order births (Jha, Kumar, Vasa, Dhingra, Thiruchelvam and Moineddin, 2006). Secondly, there are substantially more first births than subsequent births, allowing for a precise estimation of the "natural" percentage boys born in India if there are no sex selection. Finally, the results provide an indication of whether first spell length is a good indicator for fecundity.

Figures E.1, E.2, and E. 3 show the predicted percentage boys born by quarter from marriage to first birth and the associated survival functions for the lowest, middle, and highest education groups for representative women. For the first spell the representative woman is 16 years old at the beginning of the spell for the no education group, 17 years old for the middle education group, and 20 years old for the high education group. Each column represents a time period with the top panel showing urban results and the bottom panel rural results. The graphs also show the expected natural rate of boys, approximately 51.2 percent. ${ }^{6}$

[^3]The most interesting result is how close to the natural sex ratio the predicted percentage boys is for each group and for each period. As Figure E.1(d) shows, for rural women without education before 1985, who also represent the biggest group, the predicted sex ratios align almost perfectly with the expected sex ratio. For the other groups there is more volatility in the predicted percentage boys, but nowhere is it statistically significantly larger than 51.2 percent. ${ }^{7}$ Furthermore, for quarters with more substantial deviations from the natural sex ratio, the predictions are generally based on few births. In other words, it appears that the probability of having a boy is exactly the same in India as it is in other places.

For the group most likely to use sex selection, highly educated, urban women in the 1995-2006 period, the predicted percentage boys is also almost perfectly aligned with the expected percentage boys, as shown in Figure E.3(c). Hence, there is no evidence that Hindus in India use sex selection on first births. This cast serious doubts on the data used by Jha et al. (2006) and their results, as also discussed by George (2006) and Bhat (2006).

For all education groups and for all periods more than ninety percent of women had their first child within 21 quarters of being married and the proportion is increasing in education. Furthermore, 70 to 85 percent of women will had their first child within ten quarters ( 2.5 years ) of their marriage and the average time between marriage and first birth has become shorter over time. The most likely explanation for the reductions in duration and the increase in the number of women who have their first child before 21 quarters is improvements in health status. This is also consistent with the differences between education groups where more educated women are healthier and therefore more likely to conceive. There are two implications of this. First, it reinforces the need for estimating the models separately for different education levels. Secondly, it confirms that the first spell length can serve as a suitable proxy for fecundity and that Hindu women in India have their first birth very soon after marriage, even among highly educated, urban women.
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.
${ }^{7}$ The urban no education group for the 1972-1984 period show two quarters where the predicted percentage boys is just statistically significantly higher than 51.2 percent, but this is likely due to recall error not perfectly caught by the method above and the periods around those two quarters are below the natural percentage boys.

## C Comparing the Hazard and Simple Models

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. The simple model cannot estimate fertility or take into account spacing, but it is possible to compare the two methods within a spell for two outcomes: the sex ratio and the number of abortions.

For each parity, the simple model is

$$
\begin{equation*}
Y_{i}=\gamma+\alpha^{\prime} \mathbf{Z}_{\mathbf{i}}+\beta^{\prime} \mathbf{X}_{i}+\varepsilon_{i}, \tag{1}
\end{equation*}
$$

where $Y$ is a dummy variable that takes the value 1 if the child born is a boy and zero if it is a girl, and the vectors of explanatory variables, $\mathbf{Z}_{i}$ and $\mathbf{X}_{i}$, are the same as for the proposed method, except that $\mathbf{Z}_{i}$ obviously does not include the baseline hazard and that there are no time dependence. To ease presentation, the indicator for parity number is not shown. I estimate this using a logit model, although a linear probability model leads to similar results.

A major difference between the two methods is how they treat censored observations, that is women who have not yet had a given parity child. The simple model uses only observed births and the predicted sex ratio is simply what we observe for the relevant uncensored sample. The hazard model instead uses both censored and uncensored spells, which allows me to predict both how many women will have the next birth and the final sex ratio for that parity, under the assumption that the patterns of spacing and use of sex selection remain the same.

Two factors determine the differences in the predicted sex ratio and number of abortions between the two models. First, how many women have censored spells but will have a birth after the survey. Second, how much use of sex selection changes within a spell. Consider four possible combinations of the two factors. First, there are few censored birth spells and use of sex selection is constant over the spell. In that case, the predictions will be similar across the two models. Second, there are many censored birth spells and use of sex selection is constant over the spell. The pre-
dicted sex ratios will be similar, but the simple model will predict a too low number of abortions per woman and fertility will appear to be falling faster than it is. Third, there are few censored birth spells and use of sex selection changes substantially with spell length. The sex ratios again are close to each other in the two models, but the estimated use of abortions will be biased in the simple model because it cannot account for the changes in use of sex selection over the spell. Finally, there are many censored birth spells and use of sex selection changes substantially with spell length. In this case the predicted sex ratio based on the simple model will be biased. Specifically, if use of sex selection declines over the spell, the simple model will predict a higher final sex ratio than the hazard model. Similarly, the predicted number of abortions will also be biased. With sex selection and falling fertility censored spells become more and more likely, but how much this matter is an empirical question.

To determine the predicted number of abortions I use the natural sex ratio. On average, 105 boys should be born per 100 girls born, which corresponds exactly with what we observe for first births in India as seen above. Hence, if we observe $b$ boys, then we should expect to see $b \frac{100}{105}$ girls. If we observe $g$ girls the number of abortions is therefore $b \frac{100}{105}-g$. The standard approach is to use the observed sex ratio for a given parity and calculate the number of abortions using this formula. For the hazard model it is, however, possible to calculate the number of predicted abortions for each time period based on the predicted sex ratio and the number of women predicted to exit with a birth in that period using the same formula. Summing across the entire period covered by the hazard model provides a predicted total number of abortions. ${ }^{8}$

I focus on those cases where there is evidence of sex-selective abortions, the spells from parity one to two and two to three for women with eight or more years of education. ${ }^{9}$ Table C. 1 shows the number of women who enter the second and third spells, the number of births, observed in the case of the simple model and predicted in the case of the hazard model, predicted number of

[^4]abortions, and the sex ratio.
I begin with women who had a girl as their first child. For rural women in the 1985-1994 period relatively few censored women are predicted to go on to have a second birth (only 5 percent of the starting population is predicted to have a second birth and have not yet had one). In addition, there is only some evidence of sex-selective abortions. Hence, the predicted sex ratios for the simple and hazard models are close. There are slightly more abortions for the hazard model, but the abortion rates (number of abortions per birth) are fairly similar between the observed births and those predicted to have a birth later.

Urban women in the same time period follow a similar pattern with relatively few censored women predicted to move on to a second birth (around 2 percent of starting population), but the abortion rate for those additional births is 8.1 abortions per 100 births, whereas it is only 6.7 for the observed births, as we would expect from the increasing probability of having a son with increasing spell length seen in panel (b) in Figure E.8. Despite this the resulting difference in sex ratios is relatively small because of the small number of censored women predicted to exit with a birth.

Not surprisingly there are substantially more censored observations for the 1995-2006 period. For urban women about 19 percent of the starting population are predicted to have a second birth but have not yet had one and for rural women it is 23 percent. This allows for potentially larger differences in predicted sex ratios and number of abortions. Despite the large number of censored observations for rural women there is relatively little difference in the predicted sex ratio. The more interesting difference is for urban women in the 1995-2006 period. The hazard model predicts 100 more abortions by the end of the spell, but the simple model overestimates the final sex ratio. The simple model predicts 143.5 boys and the hazard model 141.3 boys per 100 girls. This is consistent with the very high probability of having a son early on in the spell followed by a somewhat lower probability for later births seen in panel (c) in Figure E.8.

The lower two panels of Table C. 1 presents the predicted outcomes for urban and rural women for the third spell, split by the sex composition of their first two children and when the spell be-

Table C.1: Comparison of Hazard and Simple Models for Women with 8 or More Years of Education

|  | Urban |  | Rural |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Simple | Hazard | Simple | Hazard |
|  | Second Birth - 1985-1994 |  |  |  |
|  | 1 girl |  |  |  |
| Starting population ${ }^{\text {a }}$ | 4,869 | 4,869 | 3,105 | 3,105 |
| Second birth ${ }^{\text {b }}$ | 3,852 | 3,926 | 2,653 | 2,795 |
| Number of abortions | 260 | 266 | 110 | 125 |
| Sex ratio (boys per 100 girls) | 120.3 | 120.7 | 114.7 | 115.1 |


|  | Second Birth - 1995-2006 |  |  |  |
| :--- | :---: | :---: | :---: | :---: |
|  | 1 girl |  |  |  |
| Starting population $^{\mathrm{a}}$ | 3,774 | 3,774 | 2,823 | 2,823 |
| Second birth $^{\mathrm{b}}$ | 1,996 | 2,704 | 1,745 | 2,407 |
| Number of abortions | 288 | 388 | 180 | 268 |
| Sex ratio (boys per 100 girls) | 143.5 | 141.3 | 131.5 | 132.2 |

Third Birth - 1985-1994

|  | 1 boy and 1 girl |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Starting population $^{\mathrm{a}}$ | 3,460 | 3,460 | 1,987 | 1,987 |  |
| Third birth $^{\mathrm{c}}$ | 1,395 | 1,558 | 1,132 | 1,336 |  |
| Number of abortions | 48 | 38 | 71 | 88 |  |
| Sex ratio (boys per 100 girls) | 109.7 | 110.4 | 119.6 | 120.2 |  |
|  | 2 girls |  |  |  |  |
| Starting population $^{\mathrm{a}}$ | 1,654 | 1,654 | 982 | 982 |  |
| Has third birth $^{\mathrm{c}}$ | 1,022 | 1,092 | 764 | 853 |  |
| Number of abortions $^{\text {Sex ratio (boys per } 100 \text { girls) }}$ | 251 | 170.0 | 266 | 97 |  |


|  | Third Birth - 1995-2006 |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | 1 boy and 1 girl |  |  |  |
| Starting population ${ }^{\text {a }}$ | 2,357 | 2,357 | 1,819 | 1,819 |
| Third birth ${ }^{\text {d }}$ | 498 | 775 | 581 | 998 |
| Number of abortions | 10 | 25 | 48 | 102 |
| Sex ratio (boys per 100 girls) | 109.8 | 112.1 | 124.6 | 129.8 |
|  | 2 girls |  |  |  |
| Starting population ${ }^{\text {a }}$ | 1,000 | 1,000 | 863 | 863 |
| Third birth ${ }^{\text {d }}$ | 417 | 600 | 461 | 661 |
| Number of abortions | 104 | 131 | 95 | 135 |
| Sex ratio (boys per 100 girls) | 165.8 | 166.0 | 161.0 | 161.0 |

[^5]gan. ${ }^{10}$ For urban women with two girls in the 1985-1994 period there are only 77 more births for the hazard model, but the difference in number of abortions is 15 with the hazard model predicting 266 abortions and the simple model 251. The resulting sex ratio is 175.3 for the hazard model versus "only" 170.0 for the simple model. In other words, it appears that the simple model underestimates the spread of the use of sex-selective abortions during the expansion of access to prenatal sex determination for this group of urban women. As discussed in the paper, use of sex selection for urban women with two girls appears to have fallen from the 1985-1994 period to the 1995-2006 period, but this is mainly an indication that sex-selective abortions are being used more during the spell to second birth for this group of women for the 1995-2006 period. The result is fewer women with two girls for their first two births and those who do are substantially less likely to have a third birth than they were before.

For rural women the main difference between the two models is for those with one boy and one girl in the 1995-2006 period. Only about $60 \%$ of the birth predicted by the hazard model are observed in the simple model, leading to substantial censoring. Furthermore, use of sex-selective abortions is predicted to be substantial for those remaining births resulting in a sex ratio of 129.8 for the hazard model and 124.6 for the simple model. As discussed in the paper, this case is the exception to the rule that sex selection does not appear to be used by families that already have one son. Presumably the use by rural women in this case is an insurance against mortality.

[^6]
## D Migration

Table D.1: Migration by Beginning of Spell

|  | Years of Education |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | None |  |  | 1-7 Years |  |  | 8 or More Years |  |  |
|  | $\begin{gathered} 1972- \\ 1984 \end{gathered}$ | $\begin{gathered} 1985- \\ 1994 \end{gathered}$ | $\begin{gathered} 1995- \\ 2006 \end{gathered}$ | $\begin{gathered} 1972- \\ 1984 \end{gathered}$ | $\begin{gathered} 1985- \\ 1994 \end{gathered}$ | $\begin{gathered} 1995- \\ 2006 \end{gathered}$ | $\begin{gathered} 1972- \\ 1984 \end{gathered}$ | $\begin{gathered} 1985- \\ 1994 \end{gathered}$ | $\begin{gathered} 1995- \\ 2006 \end{gathered}$ |
| After marriage (\%) |  |  |  |  |  |  |  |  |  |
| Did not move | 51.20 | 59.60 | 64.04 | 50.24 | 58.38 | 63.68 | 42.16 | 51.07 | 61.07 |
| Within same type | 39.76 | 32.69 | 27.99 | 34.85 | 29.94 | 25.95 | 41.38 | 34.69 | 27.35 |
| Rural to urban | 7.20 | 5.91 | 5.87 | 10.91 | 8.18 | 7.33 | 11.91 | 9.93 | 7.75 |
| Urban to rural | 1.85 | 1.80 | 2.10 | 4.00 | 3.49 | 3.04 | 4.55 | 4.30 | 3.83 |
| After first birth (\%) |  |  |  |  |  |  |  |  |  |
| Did not move | 80.66 | 86.63 | 88.78 | 75.16 | 82.93 | 88.32 | 66.60 | 76.85 | 86.33 |
| Within same type | 14.06 | 9.23 | 7.02 | 16.11 | 10.95 | 7.12 | 24.16 | 16.21 | 9.44 |
| Rural to urban | 4.25 | 3.28 | 3.23 | 6.39 | 4.33 | 3.36 | 6.63 | 4.97 | 3.04 |
| Urban to rural | 1.03 | 0.85 | 0.97 | 2.34 | 1.78 | 1.21 | 2.61 | 1.96 | 1.18 |
| After second birth (\%) |  |  |  |  |  |  |  |  |  |
| Did not move | 80.66 | 85.40 | 85.77 | 75.17 | 81.42 | 85.08 | 66.64 | 74.58 | 82.48 |
| Within same type | 14.08 | 10.12 | 8.91 | 16.15 | 11.93 | 9.33 | 24.08 | 17.59 | 11.84 |
| Rural to urban | 4.22 | 3.56 | 4.02 | 6.37 | 4.72 | 4.07 | 6.72 | 5.62 | 4.09 |
| Urban to rural | 1.04 | 0.93 | 1.30 | 2.32 | 1.94 | 1.51 | 2.56 | 2.20 | 1.58 |
| After third birth (\%) |  |  |  |  |  |  |  |  |  |
| Did not move | 80.76 | 84.22 | 84.47 | 75.19 | 81.01 | 82.24 | 68.55 | 75.25 | 82.54 |
| Within same type | 13.98 | 10.89 | 9.98 | 15.88 | 12.13 | 11.51 | 21.98 | 16.62 | 10.82 |
| Rural to urban | 4.25 | 3.91 | 4.28 | 6.57 | 4.93 | 4.74 | 7.02 | 5.89 | 4.98 |
| Urban to rural | 1.01 | 0.98 | 1.26 | 2.35 | 1.93 | 1.51 | 2.46 | 2.24 | 1.66 |

Note. "Within same type" indicates either a move from a rural area to another rural area, or a move from an urban area to another urban area.

## E All Graphs using Main Sample



Figure E.1: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with no education by quarter ( 3 month period). Predictions based on age 16 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## Urban



Figure E.2: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 17 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

Urban


Figure E.3: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 20 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## E. 1 Second Spell

First child a girl


Figure E.4: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with no education by quarter ( 3 month period). Predictions based on age 18 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First child a girl


Figure E.5: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with no education by quarter ( 3 month period).
Predictions based on age 18 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First child a girl

(a) 1972-1984 $(\mathrm{N}=1,342)$

Prob. boy (\%)


Prob. no birth yet

(d) 1972-1984 ( $\mathrm{N}=1,461)$

Prob. boy (\%)


Prob. no birth yet

(b) 1985-1994 ( $\mathrm{N}=1,969$ )


Prob. no birth yet


First child a boy
(e) 1985-1994 ( $\mathrm{N}=2,182$ )

Prob. boy (\%)


Prob. no birth yet

(c) 1995-2006 $(\mathrm{N}=1,082)$

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 $(\mathrm{N}=1,104)$

Prob. boy (\%)


Prob. no birth yet


Figure E.6: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 19 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First child a girl

(a) 1972-1984 $(\mathrm{N}=2,603)$

Prob. boy (\%)


Prob. no birth yet

(d) 1972-1984 ( $\mathrm{N}=2,803$ )

Prob. boy (\%)


Prob. no birth yet

(b) 1985-1994 ( $\mathrm{N}=4,095$ )


Prob. no birth yet


First child a boy
(e) 1985-1994 $(\mathrm{N}=4,203)$

Prob. boy (\%)


Prob. no birth yet

(c) 1995-2006 $(\mathrm{N}=2,159)$


Prob. no birth yet


Prob. boy (\%)



Figure E.7: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 19 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First child a girl


Figure E.8: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 22 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First child a girl


Figure E.9: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 22 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

## E. 2 Third Spell

First two children girls


First two children one boy and one girl


Figure E.10: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

## First two children boys



Figure E.10: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children girls


First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=5,382$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=10,473$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=3,779$ )



Figure E.11: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with no of education by quarter (3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

## First two children boys



Figure E.11: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children girls


First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=883$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 $(\mathrm{N}=1,673)$

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=848$ )

Prob. boy (\%)



Figure E.12: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First two children boys



Figure E.12: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First two children girls



First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=1,639$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=3,222$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=1,773$ )

Prob. boy (\%)



Figure E.13: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children boys


Figure E.13: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children girls


First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=1,409$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=3,460$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=2,357$ )

Prob. boy (\%)



Figure E.14: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children boys


Figure E.14: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children girls


First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=700$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=1,987$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=1,819)$

Prob. boy (\%)



Figure E.15: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First two children boys



Figure E.15: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## E. 3 Fourth Spell

## First three children girls



First three children one boy and two girls


Figure E.16: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First three children two boys and one girl

(g) 1972-1984 ( $\mathrm{N}=389$ )

Prob. boy (\%)


Prob. no birth yet

g) 1972-1984 ( $\mathrm{N}=141$ )

Prob. boy (\%)


Prob. no birth yet


First three children three boys


Prob. no birth yet

(i) 1995-2006 $(\mathrm{N}=486)$


Prob. no birth yet

) 1995-2006 ( $\mathrm{N}=155$ )
Prob. boy (\%)



Figure E.16: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First three children girls



First three children one boy and two girls
(d) 1972-1984 ( $\mathrm{N}=2,065$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=5,990$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=2,456$ )



Figure E.17: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with no of education by quarter (3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First three children two boys and one girl


Figure E.17: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with no of education by quarter (3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First three children girls



First three children one boy and two girls


Figure E.18: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First three children two boys and one girl


Figure E.18: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 1 to 7 years of education by quarter (3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First three children girls



First three children one boy and two girls


Figure E.19: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First three children two boys and one girl


Figure E.19: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First three children girls



First three children one boy and two girls
(d) 1972-1984 ( $\mathrm{N}=353$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=1,062$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=590$ )

Prob. boy (\%)


Prob. no birth yet


Figure E.20: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First three children two boys and one girl

(g) 1972-1984 ( $\mathrm{N}=87$ )


Prob. no birth yet


First three children three boys
(h) 1985-1994 ( $\mathrm{N}=681$ )


Prob. no birth yet



Prob. no birth yet

(i) 1995-2006 $(\mathrm{N}=344)$

Prob. boy (\%)


Prob. no birth yet

(i) 1995-2006 ( $\mathrm{N}=110)$

Prob. boy (\%)


Prob. no birth yet


Figure E.20: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First three children girls



First three children one boy and two girls
(d) 1972-1984 ( $\mathrm{N}=176$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 $(\mathrm{N}=664)$

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=564$ )

Prob. boy (\%)


Prob. no birth yet


Figure E.21: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First three children two boys and one girl


Figure E.21: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## F Graphs using Women Married less than 19 Years

## Urban

(a) $1972-1984(\mathrm{~N}=2,529)$

(d) 1972-1984 ( $\mathrm{N}=16,293$ )

Prob. boy (\%)


(b) 1985-1994 ( $\mathrm{N}=3,831$ )

Prob. boy (\%)


Prob. no birth yet


Rural
(e) 1985-1994 ( $\mathrm{N}=21,631$ )

Prob. boy (\%)


Prob. no birth yet

(c) 1995-2006 ( $\mathrm{N}=1,762$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=7,922$ )

Prob. boy (\%)



Figure F.1: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with no education by quarter ( 3 month period). Predictions based on age 16 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## Urban



Figure F.2: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 17 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

Urban


Figure F.3: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 20 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## F. 1 Second Spell

## First child a girl



Figure F.4: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with no education by quarter ( 3 month period). Predictions based on age 18 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First child a girl


Figure F.5: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with no education by quarter ( 3 month period). Predictions based on age 18 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First child a girl


Figure F.6: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with 1 to 7 years of education by quarter ( 3 month period).
Predictions based on age 19 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First child a girl


Figure F.7: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with 1 to 7 years of education by quarter ( 3 month period).
Predictions based on age 19 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First child a girl


Figure F.8: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with 8 or more years of education by quarter ( 3 month period).

Predictions based on age 22 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

First child a girl


Figure F.9: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with 8 or more years of education by quarter ( 3 month period).

Predictions based on age 22 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

## F. 2 Third Spell

First two children girls


First two children one boy and one girl


Figure F.10: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

## First two children boys



Figure F.10: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children girls


First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=2,803$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=8,265$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=3,661$ )



Figure F.11: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with no of education by quarter (3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

## First two children boys



Figure F.11: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children girls


First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=433$ )

Prob. boy (\%)


(e) 1985-1994 ( $\mathrm{N}=1,188$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=804$ )

Prob. boy (\%)



Figure F.12: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First two children boys



Figure F.12: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 1 to 7 years of education by quarter (3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First two children girls



First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=829$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=2,549$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=1,737$ )

Prob. boy (\%)



Figure F.13: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children boys


Figure F.13: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children girls


First two children one boy and one girl
(d) 1972-1984 $(\mathrm{N}=778)$

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=2,543$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=2,309$ )

Prob. boy (\%)



Figure F.14: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## First two children boys



Figure F.14: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children girls


First two children one boy and one girl
(d) 1972-1984 ( $\mathrm{N}=412$ )

Prob. boy (\%)


Prob. no birth yet

(e) 1985-1994 ( $\mathrm{N}=1,646$ )

Prob. boy (\%)


Prob. no birth yet

(f) 1995-2006 ( $\mathrm{N}=1,805$ )

Prob. boy (\%)



Figure F.15: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

First two children boys


Figure F.15: (Continued) Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## G Graphs using Women Dropped Because of Recall Error

1950-1980


Figure G.1: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with no education by quarter ( 3 month period). Predictions based on age 16 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.


Figure G.2: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 17 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.


Figure G.3: Predicted probability of having a boy and probability of no birth yet from time of marriage for women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 20 at marriage. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## G. 1 Second Spell

1954-1984


Figure G.4: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with no education by quarter ( 3 month period). Predictions based on age 18 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.


Figure G.5: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with no education by quarter ( 3 month period).
Predictions based on age 18 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.


Figure G.6: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 19 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.


Figure G.7: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 19 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.


Figure G.8: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 22 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.


Figure G.9: Predicted probability of having a boy and probability of no birth yet from nine months after first birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 22 at first birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

## G. 2 Third Spell

1957-1984


Figure G.10: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with no of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

1957-1984


Figure G.11: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with no of education by quarter (3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal.

N indicates the number of women in the relevant group in the underlying samples.

1958-1984


Figure G.12: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

1958-1984


Figure G.13: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 1 to 7 years of education by quarter ( 3 month period). Predictions based on age 21 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

1960-1984


Figure G.14: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for urban women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.


Figure G.15: Predicted probability of having a boy and probability of no birth yet from nine months after second birth for rural women with 8 or more years of education by quarter ( 3 month period). Predictions based on age 24 at second birth. Left column shows results prior to sex selection available, middle column before sex selection illegal and right column after sex selection illegal. N indicates the number of women in the relevant group in the underlying samples.

## References

Bhat, P N Mari, "Sex Ratio in India," Lancet, 2006, 367 (9524), 1725-1726.

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

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.


[^0]:    ${ }^{1}$ NFHS- 1 probe for each calendar birth interval that is 4 or more years. NFHS-2 asked for stillbirths, spontaneous and induced abortions and also probed for each calendar birth interval 4 or more years. NFHS-3 did not directly use birth intervals, but asked whether were there any other live births between (name of previous birth) and (name), including any children who died after birth. It did ask for births before the birth listed as first birth.
    ${ }^{2}$ As an example, imagine a family whose first child is a girl. Because of son preference this girl is followed quickly by another child. If this child is also a girl the parents will quickly try for a third child. This can lead to three children in three years and if the middle child dies and is not listed the enumerators will not probe for missing births because the interval between the first and last child is only two years.

[^1]:    ${ }^{3}$ Births in the 1970-1974 cohort, for example, took place between 18 and 23 years ago for NFHS-1, between 24 and 29 years ago for NFHS-2, and between 31 and 36 years ago for NFHS-3.

[^2]:    ${ }^{4}$ For first-borns it is between 1970-1974 and 1975-1979 for NFHS-3 and for second-borns it is between 1975-1979 and 1980-1984 for NFHS-1.
    ${ }^{5}$ The graph for second births shows a similar pattern with the likelihood of the second child being a boy going up with increasing marriage duration. The graphs for the second births and the individual survey rounds are available upon request.

[^3]:    ${ }^{6}$ 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, $b \frac{100}{105}$ girls should be born. If $g$ girls are observed the number of abortions is therefore $b \frac{100}{105}-g$ and the percent aborted

[^4]:    ${ }^{8}$ There is no difference between the two method in the predicted number of abortions for those women for whom we observe the subsequent birth.
    ${ }^{9}$ The results for the spell from parity three to four are available on request. They are not presented because the number of educated women who end up with three births is relatively small and the probability of having a fourth birth is small, which makes comparing the simple and hazard models with any precision difficult.

[^5]:    Note. The simple models are estimated using logit on uncensored observations only. The hazard models are estimated using all observations. For the second birth it covers the period from beginning of spell to 6 years ( 24 quarters) after the birth of the first child. For the third birth it covers the period from beginning of spell to 7.25 years ( 29 quarters) after the birth of the second child for the period 1985-1994 and from beginning of spell to 5.75 years ( 23 quarters) after the birth of the second child for the period 1995-2006.
    ${ }^{\text {a }}$ Number of women in period with given sex composition of prior child/children.
    ${ }^{\mathrm{b}}$ For simple model the number of women who are observed to have a second birth. For hazard model the predicted number of births that will occur between beginning of the spell and 6 years ( 24 quarters) after the birth of the first child.
    ${ }^{c}$ For simple model the number of women who are observed to have a third birth. For hazard model the predicted number of births that will occur between beginning of the spell and 7.25 years ( 29 quarters) after the birth of the second child.
    ${ }^{\mathrm{d}}$ For simple model the number of women who are observed to have a third birth. For hazard model the predicted number of births that will occur between beginning of the spell and 5.75 years ( 23 quarters) after the birth of the second child.

[^6]:    ${ }^{10}$ There is no evidence of sex-selective abortions and very few births for women with two boys and those predictions are therefore not presented.

