Son Preference, Birth Spacing, and Sex-Selective Abortions*

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

April 2017

*I am grateful to Andrew Foster and Darryl Holman for discussions about the method. I owe thanks to three anonymous referees, Shelly Lundberg, Daniel Rees, David Ribar, Hendrik Wolff, seminar participants at University of Copenhagen, University of Michigan, University of Washington, University of Aarhus, 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.

Abstract

This paper examines how birth spacing interact with the use of sex-selective abortions. I introduce a statistical method that incorporates that 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 use of sex selection leads to *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, whereas women with less education 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. 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

Spacing between births has long been used as a measure of son preference (Leung, 1988). Before prenatal sex determination became available, the only recourse for parents who wanted a son—but did not yet have one—was to have the next birth sooner. Son preference was therefore often associated with shorter average spacing after the births of girls than boys, which, in turn, was associated with worse health outcomes for girls.[1]

What has not previously been appreciated is that the introduction of prenatal sex determination fundamentally changed the relationship between son preference and birth spacing. This is because each sex-selective abortion increases the duration between observed births by approximately a year.[2] As a result, we now have a situation where families with the *strongest* son preference likely show the *longest* spacing after the birth of a daughter. But, to further complicate matters, we may still observe short spacing after the birth of daughter as a representation of son preference for families who—for one reason or another—do not use prenatal sex selection.

Spacing, by itself, can therefore no longer be used to measure son preference. Spacing is, however, still important and useful in understanding son preference, but only if combined with the likelihood of observing a boy or a girl. Short spacing after a girl is indicative of son preference if observed births are close to the natural sex ratio, whereas long spacing is indicative of son preference if observed births are male biased.

[1] See, for example, Das (1987), Rahman and DaVanzo (1993), Pong (1994), Haughton and Haughton (1996), Arnold (1997), and van Soest and Saha (2012) on shorter birth spacing after girls and Arnold, Choe and Roy (1998), Whitworth and Stephenson (2002), Rutstein (2005), and Conde-Agudelo, Rosas-Bermúdez and Kafury-Goeta (2006) on the association between shorter birth spacing and worse health outcomes for girls. Parents were also more likely to cease childbearing after the birth of a son than after a daughter (Ben-Porath and Welch, 1976; Das, 1987; Arnold, 1997; Clark, 2000; Filmer, Friedman and Schady, 2009; Altindag, 2016)

[2] This increase consists of 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). Second, an average of 6 months to next conception (Wang, Chen, Wang, Chen, Guang and French, 2003). Finally, sex determination tests are generally reliable only from 3 months of gestation onwards. These three parts add up to 12 months. The waiting time to conception does vary by woman, but even if it is very short, say one month, the additional space between births would still be 6 months per abortion.

In this paper, I introduce and apply a novel method that directly incorporates the effects of sex-selective abortions on the likelihood of a son being born *and* the duration between births. The method can be used to analyse both situations with and without prenatal sex selection. Furthermore, my proposed method allow for the duration since the last birth to affect the decision on sex selection. For example, as the duration from last birth becomes long enough, parents may reverse their decision to use prenatal sex determination, and carry the next pregnancy to term whether male or female. By examining whether sex selection decisions change with spacing, we can draw a more nuanced picture of the degree of son preference.

I apply the method to birth histories for Hindu women, using data from India's National Family and Health Surveys (NFHS), covering the period 1972 to 2006. India is a particularly interesting case. On one hand, India has seen dramatic increases in the males-to-females ratio at birth over the last three decades as access to prenatal sex determination expanded.[3] On the other hand, 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).[4] In addition, there has been substantial changes in access and legality of prenatal sex determination in India over the period covered. Abortion has been legal in India since 1971 and still is, but the first reports of on the availability of prenatal sex determination did not appear until 1982–83 (Sudha and Rajan, 1999; Bhat, 2006; Grover and Vijayvergiya, 2006). In 1994, the Central Government passed the Prenatal Diagnostic Techniques (PNDT) Act, making determining and communicating the sex of a fetus illegal.[5] Hence, the data make it possible

[3] 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 clearly not alone; both China and South Korea saw significant changes in the sex ratio at birth over the same period (Yi, Ping, Baochang, Yi, Bohua and Yongpiing, 1993; Park and Cho, 1995).

[4] 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).

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

to show how spacing between births and sex ratios have changed with the introduction of prenatal sex determination, and whether the ban affected the relationship between birth spacing and sex ratios.

There are three main findings.

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 et al., 1998; Dreze and Murthi, 2001).

2 Estimation Strategy

[TK this needs to move to later] Finally, increased reliability of access and effectiveness of contraceptive can lead to shorter spacing between births (Keyfitz, 1971; Heckman and Willis, 1976). With less reliable contraception parents choose a higher level of contraception—resulting in longer spacing—to avoid having too many children by accident. But, as contraception becomes more effective parents can more easily avoid future births, allowing them to reduce the spacing between births without having to worry about overshooting their preferred number of children. This idea may also help explain shorter spacing for better educated women than for less educated women, provided that knowledge and ability to use contraception differ across education groups (Tulasidhar, 1993; Whitworth and Stephenson, 2002).

To improve on the literature's current approaches to understanding son preference, a new empirical method must be able to capture the three main implications of the availability of sex-selective abortions combined with son preference: a higher probability that the next child is a son, longer spacing to next birth with use of sex selection, and that the use of prenatal sex determination may change between one observed birth and the next.[6] This means that we need an empirical model that can jointly estimate the association between covariates and both the sex of children born and the spacing from last birth. My proposed model is a discrete time, non-proportional competing risk hazard model with two exit states: either a boy or a girl is born.[7]

I divide a woman's reproductive life into spells that each covers the period between births (from marriage to first birth for the first spell). For each woman, $i = 1, \ldots, n$, 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.[8] 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

$$h_{ijt} = \Pr(T_i = t, J_i = j \mid T_i \geq t; \mathbf{Z}_{it}, \mathbf{X}_i), \tag{1}$$

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

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

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

[6] In addition, it should be able to capture the shorter spacing between births observed in the absence of use or availability of prenatal sex determination.

[7] 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.

[8] The time of censoring is assumed independent of the hazard rate, as is standard in the literature.

and the associated coefficients,

$$D_j(t) = \gamma_{j1} D_1 + \gamma_{j2} D_2 + \ldots + \gamma_{jT} D_T, \tag{3}$$

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.

[TK I am not very satisfied with this explanation of the problems with proportionality and the associated bias; non-proportionality is required to capture changes in use of sex selection]

In principle, specifying the model as a proportional hazard model, i.e. one where covariates simply shift the hazard rates up or down independent of spell length, is more efficient, but only provided that the proportionality assumption holds. If the proportionality assumption does not hold, however, the result is a potentially substantial bias in estimates. The problem is that a proportional hazard model does not allow covariates to have different effects at different times within a spell and therefore cannot capture differences in the shape of the hazard functions between different groups. 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. Any 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

$$\mathbf{Z}_{it} = D_j(t) \times (\mathbf{Z}_1 + Z_2 + \mathbf{Z}_1 \times Z_2), \tag{4}$$

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 (2) 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.[9] 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.

The main downside of my approach is that 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).

[9] 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 3. 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.

It is, however, straightforward to calculate the predicted probabilities of having a boy and of having a girl for each t within a spell, 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

$$P(b_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t) = \frac{\exp(D_j(t) + \alpha_{1t}'\mathbf{Z}_{kt} + \beta_1'\mathbf{X}_k)}{1 + \sum_{l=1}^{2}\exp(D_j(t) + \alpha_{lt}'\mathbf{Z}_{kt} + \beta_l'\mathbf{X}_k)}, \tag{5}$$

and the predicted probability of having a girl is

$$P(g_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t) = \frac{\exp(D_j(t) + \alpha_{2t}'\mathbf{Z}_{kt} + \beta_2'\mathbf{X}_k)}{1 + \sum_{l=2}^{2}\exp(D_j(t) + \alpha_{lt}'\mathbf{Z}_{kt} + \beta_l'\mathbf{X}_k)}. \tag{6}$$

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

$$\hat{Y}_t = \frac{P(b_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t)}{P(b_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t) + P(g_t|\mathbf{X}_k, \mathbf{Z}_{kt}, t)} \times 100, \tag{7}$$

together with the associated confidence interval for given values of explanatory variables.[10]

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.[11] For each graph, the extent to which the percent boys is statistically significantly above the natural sex ratio indicates the use of sex selection.

[TK may need to combine with and without sons to show differences in spacing over time; maybe it should also have predicted average/median spacing with some confidence intervals to test]

[10] Within each period $1 - P(b_t) - P(g_t)$ is the probability of not having a birth in period t.
[11] The parameter estimates are available on request.

The other important part of the model is the spacing between births. Spacing is captured by the survival curve, which shows the probability of not having had a birth yet by spell duration. The survival curve at time t is

$$S_t = \prod_{d=1}^{t} \left(1 - \left(P(b_d|\mathbf{X}_k, \mathbf{Z}_{kd}, d) + P(g_d|\mathbf{X}_k, \mathbf{Z}_{kd}, d)\right)\right), \tag{8}$$

or equivalently

$$S_t = \prod_{d=1}^{t} \left(\frac{1}{1 + \sum_{l=2}^{2} \exp(D_j(t) + \alpha'_{ld}\mathbf{Z}_{kd} + \beta'_l \mathbf{X}_k)} \right). \tag{9}$$

In the absence of sex selection, I expect most parities to show an "inverted s" pattern, where there initially are relatively few births, followed by a substantial number of births over a 1 to 2 year period, and then relatively few births thereafter. As sex selection becomes more widely used the associated longer spacing shows up by making the survival curve straighter, indicating that some of the births that would originally have taken place now take place later because of abortions. The use of sex selection is clearly not the only factor that can change the shape of the survival curves; factors such as the desired number of children and use of contraceptives may also shape the shape. To account for this it is best to compare survival curves for an individual parity between women who are likely to use sex selection and women who are not, for example because they already have one or more sons.

In addition to information about spacing, the survival curves also 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. Hence, although they cannot by themselves show sex selection, the survival curves are a crucial complement to the estimated sex ratios by duration.

3 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.[12] The surveys are large: NFHS-1 covered 89,777 ever-married women aged 13–49 from 88,562 households, NFHS-2 covered 90,303 ever-married women aged 15–49 from 92,486 households and NFHS-3 covered 124,385 never-married and ever-married women aged 15–49 from 109,041 households.

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.[13]

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 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 sur-

[12] A delay in the survey for Tripura means that NFHS-2 has a small number of observation collected in 2000.

[13] 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.

veys cover as long a period in the same amount of detail. The three NFHS rounds allow me to show the development in spacing and sex ratio from before sex-selective abortions were available until 2006. Third, NFHS has birth histories for a large number of women.[14] 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 establish where recall error remains a problem.

Recall error arise mainly from child mortality, when respondents are reluctant to discuss deceased children.[15] 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

[14] 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.

[15] The online Appendix contains a more thorough discussion of recall error and how I address it.