University of Massachusetts Amherst
ScholarWorks@UMass Amherst

Economics Department Working Paper Series | Economics

2009

Son Preference, Sex Selection and the Problem of Missing Women in India

Deepankar Basu
University of Massachusetts - Amherst

Follow this and additional works at: https://scholarworks.umass.edu/econ_workingpaper

 Part of the Economics Commons

Recommended Citation

Basu, Deepankar, "Son Preference, Sex Selection and the Problem of Missing Women in India" (2009). *Economics Department Working Paper Series*. 8.
https://doi.org/10.7275/1066789

This Article is brought to you for free and open access by the Economics at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Economics Department Working Paper Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

DEPARTMENT OF ECONOMICS

Working Paper

Son Preference, Sex Selection and the Problem of Missing Women in India

by

Deepankar Basu

Working Paper 2009-06

UNIVERSITY OF MASSACHUSETTS AMHERST

Son Preference, Sex Selection and the Problem of Missing Women in India

Deepankar Basu*

June 25, 2009

Abstract

This paper empirically tests for two competing explanations of the increasing sex ratio at birth (SRB) in India: hepatitis B and human intervention. Estimating a male-preferring stopping rule with data from three rounds of the National Family Health Survey in India (1992, 1998 and 2005), I find that the probability of a male birth varies significantly across birth parities. Using a novel proxy for hepatitis B in India - tribal status - I also find that hepatitis B has no impact on the probability of male birth. I conclude that human intervention explains the increasing SRB in India.

JEL Codes: J1, J7.

Keywords: son preference, sex selective abortion.

*Department of Economics, University of Massachusetts, 1012 Thompson Hall, Amherst, MA 01003, email: dbasu@econs.umass.edu. I would like to thank Robert de Jong, Trevon Logan, Alita Nandi, participants in the Micro Lunch at the Department of Economics, Ohio State University, seminar participants at the University of Massachusetts, Amherst and participants at the 2008 North American Summer Meeting of the Econometric Society in Pittsburgh for helpful comments.

1 Introduction

In the backdrop of the debate on "missing women" in South Asia and China (Sen, 1990, 1992, 2003; Oster, 2005; Das Gupta, 2005; Jha, et al., 2006; Ebenstein, 2007; Lin and Luoh, 2008; Sahni, et al., 2008; Klassen, 2008; Oster, et al., 2008), this paper tests the validity of two alternative explanations for the increasing sex ratio at birth (SRB) - defined as the number of male births per 100 female births - in India over the past two decades: biology (in the form of hepatitis B) and human intervention (in the form of sex-selective abortion and/or female infanticide). My testing procedure rests on distinguishing the effects of biology from the effects of human intervention on the SRB. To do so, I rely on the simple observation that human intervention, in the form of sex-selective abortion (or neonatal infanticide), will cause the probability of male birth to vary across birth parity (or birth order), whereas biological factors like hepatitis B, by themselves, cannot do so.[1]

Biological factors, even when they are determinants of the SRB, can be expected to operate in the same way across birth parities; hence, even if biological factors like hepatitis B did in fact affect the SRB, it would do so to the same extent at all birth parities. Thus, prevalence of hepatitis B in a population cannot account for the *variation* in the SRB across birth parities. Human intervention like sex-selective abortion, on the other hand, is a reflection of son preference and other deep gender biases in society. Thus it can be expected to vary the probability of male birth across birth parity as couples express their strong preference for male offspring in either of two ways (or both): (a) determine the sex of foetuses and induce abortion of female foetuses, or (b) practice neonatal infanticide, both conditional on the sex-composition of existing children.

To understand why son preference might lead to a significant variation in the probability of male birth across birth parities, I draw on a strand of the demography and anthropology

[1]Since there is a one-one mapping between the SRB and the probability of male/female birth, I will use the two interchangeably throughout the text.

literature that has studied this phenomenon in several parts of the world. Demographers have noted that son preference often affects fertility behaviour in countries of South Asia, South-East Asia and North Africa, taking the form of male preferring stopping rules on childbearing, also termed differential stopping behaviour (DSB) by anthropologists. DSB can be thought of in the following manner: couples continue childbearing until they attain a 'target' number of boys or hit a ceiling for the maximum number of children that they would like to have. In a nutshell, DSB makes childbearing decisions depend on the sex composition of existing children, and in conjunction with the practice of sex selective abortion of fetuses, would lead to significant variation in the probability of male birth across birth parities.[2].

To see the variation in the probability of male birth across parities, let us follow a cohort of couples who are just starting the childbearing process, assuming that each couple have positive number of children. For the first birth, let p_1 denote the probability of a male birth. For the second birth, let p_B denote the probability of a male birth conditional on the first child being a boy and let p_G stand for the probability of a male birth conditional on the first child being a girl. Then, the probability of a male birth for the second birth, p_2, is given by

$$p_2 = xp_B + (1-x)p_G,$$

where x is the fraction of families which had a boy as the first child and $(1-x)$ is the fraction of families which had a girl as the first child. If there is strong son preference in the population and technologies of sex selective abortion are widely available, then we will have

$$p_G > p_1 > p_B,$$

i.e., couples with girls as their first child will be more likely to abort a female fetus than couples with no children, while couples with a boy as the first child will be less likely to

[2]For evidence on DSB see Arnold, Choe and Roy (1998) and Larsen, Chung and Das Gupta (1998)

abort a female fetus than couples with no children.[3] Moreover, there is a particular value of x which will make the probability of male births equal for the first and second birth, i.e.,

$$\text{if} \quad x = \frac{p_G - p_1}{p_G - p_B}, \quad \text{then} \quad p_2 = p_1.$$

If x is understood as a random variable taking values in the unit interval $[0, 1]$, then with probability one, x will attain a value that is different from $(p_G - p_1)/(p_G - p_B)$; hence with probability one, we will have $p_2 \neq p_1$. Thus, we can say with certainty, that the probability of a male birth will be different between the first and the second births.

But this still does not allow us to unambiguously state whether p_2 will be greater or smaller than p_1. If $x < (p_G - p_1)/(p_G - p_B)$ then $p_2 > p_1$; if, on the other hand, $x > (p_G - p_1)/(p_G - p_B)$ then $p_2 < p_1$. This is fairly intuitive. If the proportion of families having a boy as the first child is small, then in the presence of son preference, all these families would try to attain a boy in the second birth even adopting the extreme measure of female feticide or infanticide. This would push the probability of male birth for the second birth to above its value for the first birth. If the opposite scenario were to unfold, i.e. if the proportion of families having a boy as the first child was relatively large, the probability of male birth might even fall from the first to the second birth. The crucial factor which determines whether the probability of male birth increases (or decreases) between the first and second births is the fraction of families which have had a boy as the first child. Since this fraction is difficult to predict *a priori* and since it depends on a host of factors like the exact intensity of son preference, the possibility of accidental birth, the ease of availability of sex selection technologies, etc., it is difficult to say whether the probability of male births at parity two is larger (or smaller) than its value at parity one. Nevertheless, what we can state without any ambiguity is that the two - the probability of male births at parity one

[3]This is consistent with evidence presented in Jha, et al., 2006.

and at parity two - will be unequal.

When we move to the third (and higher order) births, the picture gets more complicated. A third birth, for instance, can be preceded by four different combinations of the previous two births: two boys (BB); two girls (GG); a boy followed by a girl (BG); or, a girl followed by a boy (GB). One can plausibly argue that the probability of aborting female fetuses is higher for families with two girls than for families with a boy and a girl; and that the latter is, in turn, greater than for families with two boys. But this would still not allow us to make an unambiguous statement about the probability of male birth at parity three in comparison to the corresponding probability for parity two. This is because the probability of a male birth for the third birth parity is, just like before, a weighted average of the probability of male birth for four different cases of previous birth records, BB, GB, BG, GG, with the weights being the proportion of families of each type. Since these fractions are difficult to predict, it is not possible to state unambiguously whether the probability of a male birth would be higher or lower at parity three in comparison to the corresponding probability at parity two. The same logic would hold for births of higher parities and would prevent us from making an unambiguous prediction about how the probability of male births would vary across birth parities.

But even if we cannot make predictions about how the probability of male birth would vary across parities, i.e., whether it would monotonically increase or decrease or display a more complicated pattern, we can state with a lot of confidence that it will *vary significantly across birth parities* if there is son preference and human intervention in the form of sex selective abortion and/or neonatal female infanticide. The pattern of variation will of course depend on the exact form of human intervention and on the intensity of son preference. For instance, the variation of the probability of male birth across birth parities might display a different pattern when couples primarily practice female infanticide instead of sex-selective abortion. But in both cases - sex selective abortion and female infanticide - it will vary

significantly across birth parities. Hence, significant variation of the probability of male births across birth parities can be taken as evidence of widespread human intervention.

But probabilities of male births across birth parities are not observed; they need to be estimated. This paper contributes to the existing literature on missing women by estimating probabilities of male birth at different birth parities based on a simple model of 'son targeting' fertility behavior and evaluating whether these estimated probabilities provide evidence of gender-biased human intervention. An alternative approach works directly with sex ratios computed from the data to infer sex selection, like Jha, et al. (2006) and Ebenstein (2007). My approach has three distinct advantages over this alternative, and already existing, methodology.

First, my estimates emerge from a behavioral model of son targeting fertility behaviour which allows for sex selection, as opposed to a purely empirical approach of looking at the computed sex ratios at birth. Having an underlying model is useful because it allows us to use information about sex selection from related decision making processes like son targeting. Son preference, which leads to sex selective abortion, also affects fertility decisions relating to the 'targeting' of sons and the desired family size (i.e., the optimal number of children). Thus, jointly estimating the structural parameters pertaining to family size decisions, son targeting decisions and sex selection decisions of couples seems a better approach to follow. For instance, there is considerable evidence pointing to the existence of strong son preference in India (see, for instance, Basu and de Jong, 2009); since son targeting behaviour and family size decisions would affect sex selection (because all three are affected by underlying currents of son preference), joint estimation of the effects of covariates on all three decision making processes is more efficient. Joint estimation is, of course, only possible in the context of a behavioural model.

Joint estimation of the effect of covariates on son targeting, family size and sex selection has important implications about the sample of families that we can legitimately study.

Since the effects of son targeting fertility behaviour will only emerge when the childbearing process has run its full course, we need to focus attention on the cohort of families with completed birth histories. Studying all the families in the sample is not appropriate if we want to jointly study son targeting, sex selection and the determination of family size. That is why this paper restricts itself to the cohort of families with completed birth histories.

The second advantage of my method follows directly from the fact that I study the cohort of families with completed birth histories. Most studies that attempt to infer the presence of sex selection focus on all families in the sample during a given time period; thus, the measures of sex selection that they throw up (for instance, Jha, et al., 2006 and Ebenstein, 2007) comes from studying cross sections of families at a point in time (or during a given period of time). Since I study the cohort of families with completed birth histories instead, my estimates of the probability of male birth at different birth parities provide an alternative measure of the extent of sex selection in the population. The results in this paper suggest that focusing attention only on the sex ratios (at birth) for the whole population, as most researchers do, tends to underestimate the extent of the prevalence of sex selection in the population; one must also look at families with completed birth histories, as I do in this paper.

The third advantage of my method is that it allows one to endogenize the probability of male birth and therefore study the effect of covariates on sex selection *jointly* with related decision-making processes, like son targeting and household size decisions. This is important from a policy perspective because designing appropriate policies to tackle the problem of "missing women" and sex selective abortion would require knowledge of it's effective determinants. The effect of covariates on sex selective behaviour will be naturally provided by my approach whereas an approach based on empirically observed ratios will have to work out the effect of covariates in a separate exercise which might not account for son targeting behaviour.

The testing procedure in this paper uses Basu and de Jong (2009) who had proposed a simple framework to analyze the effect of covariates on son targeting fertility behavior. In their model, the effect of DSB on childbearing decisions was assumed to work in the following way: couples continued childbearing until they attained their 'target' number of boys or hit the ceiling for the desired number of children they wanted to have. This behavioral assumption about son targeting fertility decision making allowed them to derive the joint likelihood of any observed birth sequence and the family's desired family size. Parametrising crucial relationships underlying the joint likelihood and then estimating those parameters using MLE gave them a way to study the effect of covariates on son targeting.

While deriving the likelihood function, Basu and de Jong (2009) had assumed the probability of a male birth at every birth parity to be exogenous and fixed. For the first step of the argument in this paper, I treat the probabilities of male birth across birth parities as exogenous parameters and estimate them jointly with other parameters in the model. Then I test whether these estimated probabilities of male birth vary significantly across birth parities using a likelihood ratio (LR) test. I use birth history data provided by the three waves of National Family and Health Survey (NFHS) in India for estimation. For all the three years (1992, 1998 and 2005), I get large and positive values of the LR test statistic; hence I reject the null hypothesis of the equality of the probabilities of male birth across birth parities. I interpret this as evidence that human intervention in the form of sex-selective abortion (or neonatal female infanticide) has been prevalent in India over the last two decades.

Of course, significant variation of the estimated probability of male birth across birth parities is not, by itself, sufficient to rule out the effect of biological factors like hepatitis B. It is in principle possible for both effects, biological and social, to be present. For instance, it is conceivable that hepatitis B increases the probability of male birth at all birth parities, while sex selective abortion, acting on top of that, accounts for the significant variation across parities; hence, variation of the probabilities *per se* do not rule out the effect of hepatitis B.

To draw any conclusion about the possible link of hepatitis B to the SRB, one would need to control for its effect either directly or indirectly; that is precisely the issue that is addressed in the second step of the argument in this paper.

Controlling directly for hepatitis B was ruled out because of lack of data about the prevalence of hepatitis B in the NFHS, the source of data used for this study. Therefore I turned to the epidemiological literature on hepatitis B in India to construct a novel proxy for it and thereby indirectly control for its effect. Several studies, spanning the last two decades, provide evidence that the prevalence of hepatitis B is *significantly higher* in the scheduled tribe (henceforth ST) population in India compared to the general population. A summary of nine studies of 21 tribes spread across 8 Indian states by Murhekar and Zodpey (2005) concludes by saying: "Our study indicates that the prevalence of HBsAg among the scheduled tribes of India is much higher than in the general population" (p. 270). Based on this epidemiological finding, a dummy variable on membership to a scheduled tribe can be legitimately used as a proxy for the prevalence of hepatitis B. Note that this procedure to control for any effect of hepatitis B infection on the SRB does not depend on the route of infection; thus my procedure is valid irrespective of whether maternal or paternal hepatitis B infection affects the SRB, an issue that has been discussed in Oster (2005), Lin and Luoh (2008) and Oster, et al. (2008).

For the second step of the argument in this paper, therefore, I re-estimate the full empirical model by endogenizing the probability of male birth. I allow the probability of male birth to depend on tribal status (the proxy for hepatitis B) and also control for other covariates like age, income, education, geographical location, religion, etc. Estimation results with NFHS data for India for 1992, 1998 and 2005 show that either there is no statistically significant effect or a negative effect of the hepatitis B variation captured by tribal status on the probability of male birth. Taken together, the two steps of the argument in this paper lead to the conclusion that gender-biased human intervention, rather than the prevalence

of hepatitis B, is the leading explanation for the increasing SRB in India over the last two decades.

2 The Debate on Missing Women

In the early 1990's, Amartya Sen directed the attention of the world to what he called the problem of "missing women" (Sen, 1990; 1992). This problem refers to abnormally high population sex ratios (defined as the number of males per 100 females) in certain parts of the world, especially East, South and South-East Asia and North Africa. These skewed population sex ratios point to a large deficit of women in the adult population. By Sen's estimates, close to 100 million women were "missing". Though later attempts to refine Sen's calculations arrived at lower estimates (see, for instance, Coale, 1991), the enormity of the problem is hardly diminished because the consensus figure lies anywhere between 60 and 100 million missing women.

Sen (1992) had identified the main reason for the enormous deficit of women as the excess mortality that they faced at almost all ages despite their superior biological survival rates. Returning to the same issue a decade later, Sen (2003), noted that the gender differential in the age-specific mortality rates had declined. But this did not diminish the deficit in the number of women significantly because 'mortality inequality' had been replaced by 'natality inequality', where the latter referred to the practice of sex determination of fetuses and widespread abortion of female fetuses. Sen (1992, 2003) squarely put the blame for the problem of missing women, caused either by mortality inequality or natality inequality, on various kinds of discrimination, both overt and covert, that women face in Asian and North African societies and called for social, political and cultural measures to fight against these deep-rooted gender biases.

A recent paper (Oster, 2005) has tried to challenge the thrust of Sen's argument (see,

for instance, Dubner and Levitt, 2005) by suggesting that part of the "missing women" problem can be explained by biological factors. Drawing on the epidemiological literature on hepatitis B, Oster (2005) has argued that the prevalence of hepatitis B in a population can partly explain the skewed population sex ratios. This, according to her, is because hepatitis B infection among women seems to increase the sex ratio at birth (SRB), i.e., carriers of the hepatitis B virus give birth to more male offspring than non-carriers. In her calibration exercise the prevalence of hepatitis B can account for, on the higher end, as much as 70 percent of the missing women in China; on the lower end, she claims to be able to explain about 20 percent of the problem for India. Several countries, in Oster's (2005) analysis, lie between these two extremes.

Oster's (2005) claim, it is obvious, rests crucially on the (causal) link between hepatitis B infection rates in a population and the SRB. Several facts suggest caution in using or interpreting this link. First, there is no currently available biological (or other scientific) explanation for this supposed link (p. 1180-81, Oster, 2005). As pointed out in Lin and Luoh (2008), some researchers had hypothesized that HBV(+) women might have a higher probability of spontaneously aborting female foetuses but Lividas, et al. (1979) found no evidence for this hypothesis. So interpreting the statistical link, in both individual level and cross-country data, as a *causal* link is at best atheoretical.

Second, a recent study (Lin and Luoh, 2008) attempting to *directly* examine the link between hepatitis B infection of mothers and the sex ratio of their offspring came to a decisive negative conclusion. Using a large micro-level dataset from the National Hepatitis B Immunization Program in Taiwan the authors find that the marginal probability of a positively infected (i.e., an HBsAg+) mother having a male birth is only 0.0025. Hence the study concludes that "the effect of HBV *mothers* on the sex ratio among their offspring is minimal, and hence can account for only a small portion of the case of missing women" (Lin and Luoh, 2008:2; emphasis added). The emphasis on maternal infection was in response

to Oster (2005), who had explicitly claimed that the link from the hepatitis B virus to the offspring sex ratio runs through the mother.[4]

Oster, et al. (2008), revisiting the earlier work of Oster (2005), find that hepatitis B infection does not have any statistically significant impact on the SRB. Since Lin and Luoh (2008), responding to Oster (2005), had tested the effect of maternal hepatitis infection on the SRB, that left open the possibility that the paternal infection might be driving the results in Oster (2005). To test the effect of both maternal and paternal carrier status, Oster, et al. (2008) collect data for about 67000 individuals participating in a larger prospective cohort study of liver cancer and the effect of hepatitis B on general health status in China. The results in Oster, et al. (2008) show that parental (i.e., both maternal and paternal) hepatitis B carrier status has no effect on the SRB in China, overturning the original conclusions in Oster (2005).

There are at least two more reasons which cast doubt on the alleged link of hepatitis B infection and the SRB. First, Sub-Saharan Africa is a region that is well-known for high hepatitis B infection rates, compared to world averages, but which has no female deficit in striking contrast to, for instance, Asia. Second, several studies have documented considerable *variation* in the SRB when measured at different birth parities (Lin and Luoh, 2007), or when measured for different sex-composition of existing children (Ebenstein, 2007), or when measured for different geographical locations (Hesketh and Wei Xing, 2006), or when measured across time (Das Gupta, 2005).

Though variation of the SRB across birth parities or geographical location or time does not *per se* rule out the role of hepatitis B in the determination of the SRB, it certainly points to other important factors at work, especially human intervention through sex selective abor-

[4] "There is evidence that *women* who are carriers of hepatitis B give birth to a higher ratio of boys to girls than non-carriers. Since many of the countries with missing women also have a relatively high prevalence of hepatitis B carriers, the naturally occurring higher sex ratio at birth could produce a higher population sex ratio even in the absence of excess female mortality" (p. 1164, Oster, 2005; emphasis added).

tion and/or female infanticide. Responding to Das Gupta's (2005) critique that the hepatitis B explanation for missing women cannot be valid for China because sex ratios change over time and across families, Oster (2006) correctly distinguishes the marginal from the average effect. While variation in the SRB indicates marginal effects of human intervention, Oster's (2005) argument was about the average effect of hepatitis B on the SRB. Hence, evidence of variation in the SRB does not constitute a valid critique of Oster's (2005) position, a point that is addressed by the empirical strategy proposed in this paper.

The evidence on variation in the SRB is of course well known and convincing in several cases. For instance, drawing on a strand of the demography literature, Ebenstein (2007) provides a critique of the hepatitis B explanation for missing women in China by marshaling evidence of variation in the SRB. Using Census data, he computes the SRB for different birth parities and different combinations of already-existing children.[5] He finds that whereas the SRB for the first birth is close to what is observed in Western Europe or Sub-Saharan Africa, the SRB increases steeply for later births. Looking at the male fraction of next birth at different birth parities gives even more striking results. At the second birth parity, he finds that the male fraction of next birth for families with one daughter is higher than for families with one son; at the third birth parity, he finds that the male fraction of next birth is highest for families with two daughters, which is lower for families with a son and a daughter and even lower for families with two sons. Since the SRB for the first birth is statistically indistinguishable from the 'normal' figure for human populations, this is evidence against the claim that hepatitis B affects the SRB in China; and, since there is considerable variation across parities and conditional on the sex composition of existing children, this shows that human intervention is important. Ebenstein (2007), therefore, concludes that it is human intervention in the form of sex-selective abortion, rather than biological factors like hepatitis

[5]Demographers interested in studying the effects of son preference have typically looked at variations in the SRB over birth parities, over the sex composition of existing children, over geographical location, etc. For a recent study on similar issues see Hesketh and Wei Xing, 2006 and the references therein.

B, that seem to be more important in China.

This paper extends the critical analysis of the possible link between hepatitis B and the SRB to the case of India using a novel empirical strategy. Using data from three rounds of the NFHS, I address two related questions: (a) whether there is evidence of gender-biased human intervention, and (b) whether hepatitis B infection rates have any significant impact on the SRB.

3 The Empirical Strategy

The argument in this paper proceeds in two steps. In the first step, the presence of sex-selective abortion (or female infanticide) is established by presenting evidence of statistically significant variation of probabilities of male birth across birth parities. In the second step, the effect of hepatitis B on the SRB is studied with the use of a novel proxy for hepatitis B: tribal status. Taken together, the two steps of the argument lead to the conclusion that human intervention rather than biological factors explain the movement in the SRB in India.

My basic empirical strategy, in both steps of the argument, is to estimate, following Basu and de Jong (2009), an empirical model of son targeting fertility behaviour, but now allowing for the presence of gender-biased human intervention which can alter the probability of male/female births across parities. I assume that couples continue childbearing until they attain their target for male children or when they hit the ceiling for the maximum number of children. The empirical model consists in deriving the joint likelihood of observing a family's completed birth sequence and maximum size (i.e., total number of children). If S_i denotes the *completed* birth sequence for the i^{th} family where the sequence only includes children that are currently alive, while N_i denote the maximum number of children that family i would *like* to have, then the joint probability of observing a particular birth sequence, S_i,

and desired family size, N_i, would give us the log likelihood function as

$$l = \ln(L) = \sum_{i=1}^{n} \ln(P(S_i, N_i)), \tag{1}$$

where

$$\begin{aligned} P(S_i, N_i) = {} & P(S_i|N_i, T_i = 0)\frac{\exp(-\exp(Z_i'\gamma))[\exp(Z_i'\gamma)]^{N_i}}{N_i!}\Phi(-X_i'\beta) \\ & + \sum_{k_i=1}^{N_i-1} P(S_i|N_i, k_i, T_i = 1)P(k_i|N_i, T_i = 1) \\ & \times \frac{\exp(-\exp(Z_i'\gamma + \alpha))[\exp(Z_i'\gamma + \alpha)]^{N_i}}{N_i!}[1 - \Phi(-X_i'\beta)] \end{aligned} \tag{2}$$

and n is the number of families in the sample.[6] Maximizing this function over the parameter space gives me estimates of the relevant parameters, which includes the probabilities of male birth at different birth parities.

For the first step of the argument, I allow for gender-biased human intervention by letting the probability of male birth birth vary across birth parities and treating these as exogenous parameters in the model. The null hypothesis of the equality of these estimated probabilities is then tested through a LR test to infer the presence or absence of human intervention. It is understood that a large value of the LR statistic implies rejection of the null hypothesis and thus provides evidence for the presence of human intervention.

In the second step of the argument, I test if hepatitis B infection rates have any effect on the probability of male birth. To do so I endogenize the probability of male birth while forcing it to be equal across birth parities. Restricting the probability of male birth to be equal across birth parities follows from the understanding that biological factors like hepatitis B carrier status cannot lead to variation in the SRB across parities. Since, in this step of the argument, I am interested in studying the possible effect of hepatitis B infection on the

[6]For details of the derivation of the log likelihood function see the Appendix.

SRB, it is legitimate to force the probabilities of male birth to be equal across parities.[7] The crucial variable determining the probability of male birth is, of course, tribal status (which acts as a proxy for the prevalence of hepatitis B); other covariates like age, income, education, geographical location are included as additional controls. A statistically insignificant value of the coefficient on tribal status is interpreted as evidence against the claim that hepatitis B affects the SRB.

4 Evidence of Sex-Selective Abortion

In this section I present results of estimating the probabilities of male births across birth parities with three waves of NFHS data from India: 1992, 1998 and 2005. There are several things to note about how the NFHS data is used for estimating the empirical model in this paper. First, I restrict myself to children who were alive at the time of the interviews; this is motivated by the observation that son targeting behaviour responds to the sex composition of existing children and not to all children ever born. Second, from all the respondents in the data, I pick up families which can be reasonably assumed to have completed their fertility history.[8] Since my empirical model jointly estimates decision making relating to son targeting and sex selection, I need to restrict the sample to the cohort of families with completed birth histories. Third, I do not take account of the fact that some of the births might be accidental.[9] Fourth, for families which have an actual number of children greater

[7] One could, in principle, endogenize the probability of male birth *and* allow it to vary across birth parities. This would, of course, increase the computational burden substantially. One way to avoid the additional computational burden would be to adopt a two step estimation procedure. The first step would be a probit regression for each birth parity and the second step would use estimated probabilities in the maximum likelihood estimation and correct standard errors for the two step procedure.

[8] Respondents are asked about their "fertility preferences" and future plans about childbearing through the following question: would you like to have another child or would your prefer not to have any more children? I interpret responses of "no more", "sterilized" and "declared infecund" as indicating completion of the childbearing process.

[9] For instance, a family which has stopped after a boy and a girl (in that order) might have been targeting one boy and have had the second child accidentally. If this family reports a desired maximum of 2, in my model, this family would come out as a non-targeter whereas in fact it was targeting. Given the fact that

than what they report as their 'desired maximum', I recode their desired maximum as the actual number of children alive.

One possible shortcoming of my approach should be noted immediately. In the NFHS questionnaire, there are questions both about the desired number of sons and the desired total number of children. While it is true that answers to both these questions are possibly unreliable or untrue, I need to assume that at least one of these is true. In writing the likelihood, I assume that the desired total number of children is observed (or truly revealed) whereas the desired number of sons is either not observed or not truly revealed.[10] This is motivated by the understanding that information on (desired) total number of children might be more reliable than information on the desired son target because couples might be less willing to reveal their true son preference which is considered to be more sensitive information.

In presenting the results below I focus on two aspects. First, I present estimates of the probability of male birth at different birth parities;[11] these estimates emerge from estimating the unrestricted model, i.e., the model that allows the probabilities to vary across birth parities. Second, I test the null hypothesis that all these estimated probabilities are equal. To do so, I re-estimate the model under the null (the restricted model) i.e., the model with the restriction that the probability of male birth is the same at all birth parities; then I use the likelihood ratio (LR) statistic to test the null of equality of birth probabilities.

This argument would lose some of its force if there were biological factors which led to variation in the probabilities of male births across birth parities. Demographers and

contraception use has been rapidly increasing in India though, the proportion of accidental births can be expected to fall over time and thus not affect the results significantly.

[10]In the NFHS questionnaire, information about the desired maximum number of children is elicited by two consecutive questions. All ever-married women who had any living children were asked: "If you could go back to the time you did not have any children and could choose exactly the number of children to have in your whole life, how many would that be?" To those who did not have any living children were asked: "If you could choose exactly the number of children to have in your whole life, how many would that be?"

[11]For the estimation I consider the first through sixth birth parities. R code for estimation is available from the author upon request.

population biologists have analyzed various biological, environmental and social factors as possible determinants of the sex ratio at birth (SRB) in human populations. Though birth order (or parity) has figured in many such studies as a possible determinant of the SRB, the current consensus seems to be that birth order does not have any biological impact on the SRB (Ulizzi and Zonta, 1995; Hesketh and Xing, 2006).

Estimation results about the probabilities of male birth are presented in Table 1, Table 2 and Table 3 for years 1992, 1998 and 2005 respectively; Figure 1 visually summarizes the information in these tables about the probabilities of male birth across birth parities. The first thing to notice about the estimated probabilities for all the three time periods is that they are all significantly higher than the 'normal' figure of 0.514;[12] this suggests that either sex-selective abortion or hepatitis B or both are in operation in India right from the first birth parity.

This finding is important because it is different from Ebenstein's (2007) result about China to the effect that the SRB for the first birth is statistically indistinguishable from the 'normal' value of 0.514. The high values of the estimated probabilities for the first birth parity in the case of India, therefore, rules out using Ebenstein's (2007) argument for finding evidence of human intervention; since the estimated probability of male birth is statistically significantly greater than the 'normal' value of 0.514 at all parities, we cannot rule out the effect of hepatitis B at this stage of the argument. Note also that my finding is in line with the data about India reported in Jha, et al. (2006). That study used data for about 1.1 million births reported in the Special Fertility and Mortality Survey of February 1998. The data provided in Table 1 in Jha, et al. (2006) show that the fraction of male births at the first birth parity was 0.54383, at the second parity was 0.5316 and at the third parity was 0.5348; thus all the figures, including the one for the first birth, are significantly greater than the 'normal' value of 0.514, exactly in tune with the results in this paper. Note, in passing,

[12]The t-statistics in the tables arise from the test of the probabilities being equal to 0.514.

that a probability of male birth of around 0.53 implies about 113 boys for every 100 girls; since the normal ratio is roughly 105 boys for every 100 boys, this implies about 8 missing girls for every 100 boys, an extremely high value even by contemporary standards. For a country like India, this would translate to about 80 million missing women.

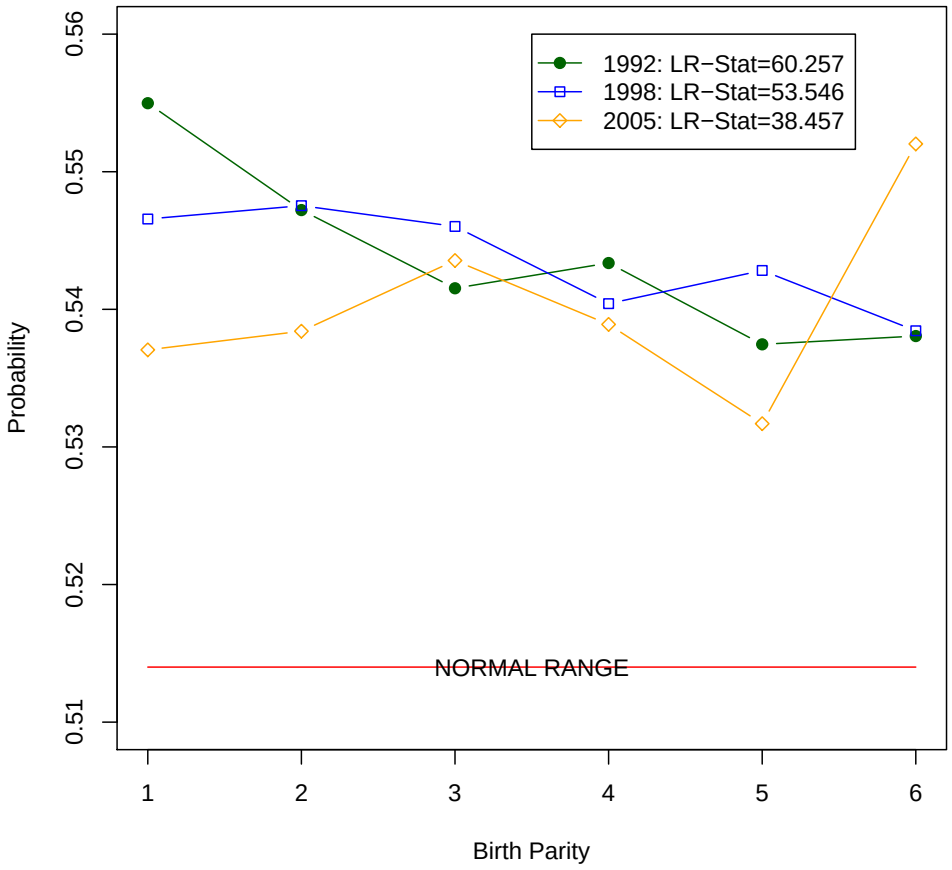

Figure 1: Probabilities of Male Birth

The second thing to notice is that the LR statistic, which is distributed as a χ^2 random variable with 5 degrees of freedom, is significantly greater than zero for all years. Recall that the LR statistic relates to the test of the null hypothesis that the probability of male birth is equal across parities; the fact that the LR statistic is large for each of the three

years suggest that the null can be rejected in each of the three years with high degrees of confidence. Thus, evidence of human intervention through sex-selective abortion or neonatal infanticide is strong for all years.

Probabilities of male birth for different birth parities have been plotted for all the three time periods in Figure 1. The fact that these probabilities vary considerably across birth parities strikes one immediately. This is more or less what one would expect to find given the presence of son preference and gender-biased human intervention. As pointed out earlier, there is no reason to expect any monotonic pattern in the probabilities of male birth across parities; the only thing one can reasonably expect to find is significant variation across parities.

Though more research is needed into this issue, my current hypothesis for explaining the shape of the probability plots in Figure 1 is that there has been a gradual change in the behaviour of the population: couples seem to have gradually shifted away from the practice of female infanticide and towards sex selective abortion. Several facts point in this direction.

First, ultra-sound technology for the determination of the sex of the fetus became widely available in India only from the late 1980s. Most of the women who had been surveyed during the 1992 Demographic and Health Survey (DHS) must have finished childbearing before the sex selection technology became readily available. Now, a case of female infanticide would always mean that *that* birth would go unreported, thereby increasing the fraction of reported male birth. Thus, high probabilities of male births (for 1992) would tend to suggest that female infanticide was probably being widely practiced because technologies for sex selective abortion were still not easily available.

Second, the shape of the probability plot for 1992, which has a downward trend, might also provide some information in this regard. It seems that sex-selective abortion and neonatal infanticide entail very different psychological costs for the couple, especially for the mother. Whereas abortion in the early stages of pregnancy might be stressful, neonatal infanticide

is considerably more so. The psychological cost of neonatal infanticide, moreover, might increase sharply with the presence of children in the family. The downward trend in the birth probabilities (for 1992) might then be explained by the fact that the existence of children reduces the probability of parents accepting neonatal infanticide as a feasible option.

Third, the fact that the shape of the probability plots change from 1992 to 1998 to 2005 suggests that the there is some change in the underlying behaviour of couples. The shape for 2005, where the probabilities increase steeply in the initial parities, is more or less in line with what one would expect when sex selection technologies are widely available.

5 Does Biology Have a Role to Play?

The fact that the estimated probabilities of male birth vary significantly across birth parities provides strong evidence of human intervention in the form of sex-selective abortion and possibly also neonatal infanticide. But this, by itself, does not rule out a role for biology. As seen in Figure 1, all points on the probability plot are significantly greater than 0.514. Hence it is conceivable that *both* biology and human intervention have a role to play.

Since NFHS data does not provide any information about hepatitis B, I need to look elsewhere for information on hepatitis B. A suitable proxy for hepatitis B is suggested by the epidemiological literature on the prevalence of hepatitis B in India. Murhekar and Zopdey (2005) summarize a substantial body of epidemiological research into the incidence of hepatitis B among tribal populations in India[13]. Bringing together nine studies about 21 tribes across 8 states in India, they conclude that the prevalence of hepatitis B is significantly higher in the tribal population than the general population. They compute the average hepatitis B prevalence rate for the tribal population as 10.15% whereas the corresponding

[13]The indigenous population of India are said to belong to Scheduled Tribes; they constitute about 10 percent of the population, are geographically isolated and sometimes live in semi-egalitarian communities. Juridically, Article 342 of the Indian Constitution prescribes procedures to be followed in the matter of specification of Scheduled Tribes.

prevalence rate for the whole population in India is about 3%.[14] I use this observation to construct a proxy for the prevalence of hepatitis B by constructing a dummy variable for whether the respondent is a tribal or not; I use this dummy to capture the possible effect of hepatitis B on the probability of male birth. Note that this empirical strategy is valid irrespective of whether hepatitis B infection affects the sex ratio through maternal or paternal infection.

I return to my full empirical model of son targeting fertility behaviour and re-estimate the model by endogenizing the probability of male birth, allowing tribal status to be one of its determinants; for this exercise I restrict myself to families with completed fertility histories, using data for the years 1992, 1998 and 2005. I restrict myself to families with completed birth histories because my empirical model of son targeting fertility behavior applies only to families which have completed their childbearing process; the likelihood function underlying my empirical model does not apply to families still within the fertility cycle.

5.1 Results from MLE of the Full Model

To incorporate the effect of hepatitis B on the probability of a male birth I add an equation to the empirical model of the male-preferring stopping rule in (1). This additional equation determines the probability of male birth for family i as a function of covariates, the most crucial among them being the tribal status dummy, which captures the possible effect of hepatitis B carrier status on the SRB. As additional controls, I use standard covariates: age of the mother, years of formal education of the mother, a dummy variable on whether the mother works outside the household for a wage, a dummy variable on whether the household resides in a rural area, two income dummies[15], a dummy on whether the family belongs

[14] See `http://www.natap.org/2005/HBV/101305_02.htm`

[15] Since there are no income variables in DHS, I constructed income/class dummies in the following manner: a respondent is designated *rich* if she owns a car/truck and has electricity connection in her house; a respondent is designated *middle* if she has electricity in her house but does not own a car/truck; and a respondent is designated poor if she neither owns a car/truck nor has an electricity connection to her house.

to the largest religious community (Hindu), a dummy on whether the household is a joint family. Thus, the equation determining probability of male birth for family i is:

$$p_i = \Phi(w_i'\delta) \tag{3}$$

where w_i are the covariates determining the probability of male birth, δ is the vector of corresponding parameters and $\Phi()$ denotes the standard normal distribution function.

Results of the estimation for 1992, 1998 and 2005 are presented in Table 4, Table 5 and Table 6 respectively. Note that the coefficient on tribal status is statistically insignificantly in 1992 and 2005, and is significantly *negative* in 1998. This demonstrates either that the prevalence of hepatitis B does not have any significant impact on the SRB or that it has a negative impact in India.[16] In either case the positive association between prevalence of hepatitis B and the SRB, whereby hepatitis B infection rates is supposed to increase the SRB, is not observed in the data for India. In Figure 2, Figure 3 and Figure 4, I present the distribution of the estimated probability of male birth for families in the sample. The median of the distribution is shifted far to the right of the 'normal' value of 0.514 for all the three years, giving indication of massive gender-biased human intervention over the last few decades.

6 Conclusion

The problem of missing women in parts of Asia and North Africa, whereby social inequalities and discrimination outweigh women's natural survival advantage and give rise to a large deficit of women in the population, has been highlighted by Sen (1990, 1992, 2003), among others. Oster (2005) challenged the thrust of Sen's argument by proposing a biological

[16] A significantly negative value of the parameter on the proxy for hepatitis B suggest that, contrary to Oster's (2005) contention, the prevalence of hepatitis B in India *increases* the probability of a *female* birth.

mechanism to explain at least part of the missing women problem. She suggested that the prevalence of hepatitis B in a population skews the SRB in favour of male offspring and could thus explain part of the female deficit in the adult population.

There is of course a different mechanism which can also affect the SRB: human intervention in the form of sex-selective abortion and/or neonatal infanticide. A cultural milieu marked by strong son preference has combined with the gradual spread of ultra-sound technology in India over the last two decades to markedly increase the prevalence and acceptance of sex-selective abortion.[17] Hence, there are two competing hypotheses about the increasing SRB in India over the last two decades: biology and sex-selective abortion.

In this paper I empirically test for these alternative hypotheses in two steps. In the first step I estimate a simple model of son targeting fertility behaviour and find that the probability of male birth varies significantly across birth parities; I interpret this as evidence of human intervention in the form of sex-selective abortion (or neonatal infanticide) since prevalence of hepatitis B by itself cannot account for such variation across birth parities. In the second step, I use a novel proxy for the prevalence of hepatitis B - membership in a scheduled tribe - to directly estimate the effect of hepatitis B on the probability of male birth. I find no statistically significant effect for 1992 and 2005, and a negative impact for 1998. Based on these two steps I conclude that biological factors, like the prevalence of hepatitis B, have no effect on the SRB in India; rather, it is direct human intervention in the form of sex-selective abortion which is more important.

Appendix

To derive the log likelihood function for the empirical model, I follow Basu and de Jong (2009). I will use the following notation: let S_i denote the *completed* birth sequence for

[17] For recent reporting on this issue, see the detailed coverage provided in BBC at the following website: `http://news.bbc.co.uk/2/hi/south_asia/4592890.stm`

the i^{th} family, with only living children being included in the completed birth sequence; for instance, if S_i was BBG (where B stands for boy and G for girl), then it would mean that this family has three children currently alive, the eldest being a boy, the next one a boy and the youngest a girl. Let N_i denote the maximum number of children that family i would *like* to have; let k_i denote the 'target' number of boys for family i. To estimate the probabilities of male births at different birth parities jointly with the effect of covariates on targeting and fertility behaviour, I will calculate the joint likelihood of observing a given birth sequence (S_i) *and* desired maximum number of children (N_i), parametrize the joint probability appropriately and then maximize it over the parameter space; in other words, I will compute a parametric expression for $P(S_i, N_i)$ where $P(.)$ denotes probability.

To proceed, let T_i be a dichotomous *unobservable* variable which indicates whether family i targets sons or not. $T_i = 0$ means that the family does not target; and $T_i = 1$ implies that family i is a son targeter. Finally, let T_i be determined by a vector of covariates, X_i, in the following manner:

$$T_i = \begin{cases} 0 & \text{if} \quad X_i'\beta + \varepsilon_i \leq 0 \\ 1 & \text{if} \quad X_i'\beta + \varepsilon_i > 0 \end{cases} \tag{4}$$

where X_i is a $(k \times 1)$ vector of co-variates which determine whether a particular family 'targets' sons or not, β is a $(k \times 1)$ vector of parameters to be estimated and $\varepsilon_i \sim N(0, 1)$ captures the unobservable, omitted, stochastic factors that affect son targeting behaviour.

To get the joint likelihood for the observed birth sequence and maximum number of

children for the i^{th} family, note that

$$
\begin{aligned}
&P(S_i, N_i) = P(S_i, N_i|T_i = 0)P(T_i = 0) + P(S_i, N_i|T_i = 1)P(T_i = 1) \\
&= P(S_i|N_i, T_i = 0)P(N_i|T_i = 0)P(T_i = 0) \\
&\quad + P(S_i|N_i, T_i = 1)P(N_i|T_i = 1)P(T_i = 1) \\
&= P(S_i|N_i, T_i = 0)P(N_i|T_i = 0)\Phi(-X_i'\beta) \\
&\quad + P(S_i|N_i, T_i = 1)P(N_i|T_i = 1)[1 - \Phi(-X_i'\beta)] \qquad (5)
\end{aligned}
$$

where $\Phi(.)$ denotes the standard normal cdf. Note that, because of DSB, when a family does not target sons, the effective stopping rule for childbirth becomes the maximum number of children that the family desires to have, N_i; hence the probability of observing S_i given N_i when the family does not target sons is

$$
P(S_i|N_i, T_i = 0) = I(n(S_i) = N_i) \prod_{j=1}^{m_i} (q_j^{1-B_j}) \times (1 - q_j)^{B_j}, \qquad (6)
$$

where $I(.)$ denotes the indicator function and $n(S_i)$ denotes the number of children in birth sequence S_i, $B_j = 0$ if the j^{th} child is a boy and $B_j = 1$ if the child is a girl, m_i is the number of children alive in family i, and q_j is the probability of a male birth at birth parity j (which, at this stage, I ssume to be constant for all families in the population). The indicator function in the product is meant to rule out the possibility that a family which stops childbearing before hitting the desired maximum number of children could be a non-targeter. Any family which stops childbearing before hitting the ceiling, N_i, has to be a son targeter in my model. In the special case when the probability of male birth does not change across birth parities (denoted, in this case, by q) and b_i is the number of boys in family i, g_i is the number of

girls in the same family, we will have:

$$P(S_i|N_i, T_i = 0) = I(n(S_i) = N_i)(q^{b_i}) \times (1-q)^{g_i}. \tag{7}$$

This gives us the first term in (5). The second term in (5) comes from son targeters; so I need to compute $P(S_i|N_i, T_i = 1)$. Note that when a family targets sons, it's target, k_i, can range anywhere from 1 to $N_i - 1$; targeting $k_i = N_i$ sons with a ceiling for the desired maximum number of children at N_i is equivalent to not targeting. Since I cannot observe k_i (the target number of sons for a family), I condition on k_i and then integrate it out as follows:

$$P(S_i|N_i, T_i = 1) = \sum_{k_i=1}^{N_i-1} P(S_i|N_i, k_i, T_i = 1)P(k_i|N_i, T_i = 1) \tag{8}$$

where $P(S_i|N_i, k_i, T_i = 1)$ is the probability of observing a birth sequence S_i given N_i, k_i and $T_i = 1$ (son targeting); $P(k_i|N_i, T_i = 1)$ is the probability of targeting k_i sons given that the desired maximum number of children is N_i. Since these targeting (conditional) probabilities cannot be observed, in my model, I treat them as parameters and estimate them jointly with other parameters. Using (8), therefore, (5) becomes:

$$\begin{aligned} P(S_i, N_i) = {} & P(S_i|N_i, T_i = 0)P(N_i|T_i = 0)\Phi(-X_i'\beta) \\ & + [1 - \Phi(-X_i'\beta)] \sum_{k_i=1}^{N_i-1} P(S_i|N_i, k_i, T_i = 1)P(k_i|N_i, T_i = 1)P(N_i|T_i = 1). \end{aligned}$$

Next, I assume that N_i conditional on T_i is distributed as a Poisson random variable with conditional mean given by λ_i. I try to capture two crucial facts with a Poisson distribution: first, that N_i conditional on T_i is a count variable; and second that there is interaction between the decision of son targeting and the desired maximum number of children (the

interaction term appears in the expression for the conditional mean, λ_i).

$$P(N_i|T_i) = \frac{\exp(-\lambda_i)[\lambda_i]^{N_i}}{N_i!},$$

where

$$\lambda_i = \exp(Z_i'\gamma + \alpha T_i). \tag{9}$$

Note that in the above expression α captures the effect of 'son targeting' on the total fertility rate and Z_i is a set of covariates which affects desired family size. Moreover, since T_i is a dichotomous variable, I have

$$P(N_i|T_i = 0) = \frac{\exp(-\exp(Z_i'\gamma))[\exp(Z_i'\gamma)]^{N_i}}{N_i!} \tag{10}$$

and

$$P(N_i|T_i = 1) = \frac{\exp(-\exp(Z_i'\gamma + \alpha))[\exp(Z_i'\gamma + \alpha)]^{N_i}}{N_i!} \tag{11}$$

Using (10), (11) and (8), I can write (5) as:

$$\begin{aligned} P(S_i, N_i) = {} & P(S_i|N_i, T_i = 0)\frac{\exp(-\exp(Z_i'\gamma))[\exp(Z_i'\gamma)]^{N_i}}{N_i!}\Phi(-X_i'\beta) \\ & + \sum_{k_i=1}^{N_i-1} P(S_i|N_i, k_i, T_i = 1)P(k_i|N_i, T_i = 1) \\ & \times \frac{\exp(-\exp(Z_i'\gamma + \alpha))[\exp(Z_i'\gamma + \alpha)]^{N_i}}{N_i!}[1 - \Phi(-X_i'\beta)] \end{aligned} \tag{12}$$

The log-likelihood for the observed sample, then, becomes:

$$l = \ln(L) = \sum_{i=1}^{n} \ln(P(S_i, N_i)) \tag{13}$$

where n is the number of families in the sample and $P(S_i, N_i)$ is substituted from (12). Maximizing this function over the parameter space gives me estimates of the relevant parameters (which includes the probabilities of male birth at different birth parities).

References

Arnold F, Choe M. K. and T. K. Roy. 1998. "Son Preference, the Family-building Process and Child Mortality in India." *Population Studies*, 52: 301-315.

Basu, D. and R. de Jong. 2009. "Son Targeting Fertility Behaviour: Some Consequences and Determinants." *Demography*, (forthcoming).

Coale, A. J. 1991. "Excess Female Mortality and the Balance of the Sexes in the Population: An Estimate of the Number of 'Missing Females."' *Population and Development Review*, 17: 517-23.

Das Gupta, M. 2005. "Explaining Asia's "Missing Women": A New Look at the Data." *Population and Development Review*, 31(3): 529-35.

Dubner, S. J. and S. D. Levitt. 2005. "The Search for 100 Million Missing Women: An Economics Detective Story." `http://www.slate.com/id/2119402/`, (accessed February 10, 2009).

Ebenstein, A. Y. 2007. "Fertility Choices and Sex Selection in Asia: Analysis and Policy." http://www.demog.berkeley.edu/∼ebenstei/ (accessed September 19, 2007).

Hesketh, T. and Z. Wei Xing. 2006. "Abnormal Sex Ratios in Human Populations: Causes and Consequences." *Proceedings of the National Academy of Sciences of the United States of America*, 103: 13271-13275.

Jha, P., Kumar R., Vasa, P., Dhingra, N., Thiruchelvam, D., and R. Moineddin. 2006. "Low Female-to-Male Sex Ratio of Children Born in India: National Survey of 1.1 Million Households." *Lancet*, 267: 211-18.

Klasen, S. 2008 "Missing Women in South Asia and China: Biology or Discrimination?" http://www.voxeu.org/index.php?q=node/1583, (accessed February 09, 2009).

Larsen U., Chung W. and M. Das Gupta. 1998. "Fertility and Son Preference in Korea." *Population Studies*, 52: 317-325.

Lin, M-J. and M-C. Luoh. 2008. "Can Hepatitis B Mothers Account for the Number of Missing Women? Evidence from Three Million Newborn in Taiwan." *American Economic Review*, (forthcoming).

Lividas, D., D.A. Koutras, J. Economidou, J.E. Hesser and S. Hadziyannis. 1979. "Fertility and Sex Ratio of Offspring of Female HBsAg Carriers," *Journal of the Royal Society of Medicine*, 72(7): 509-512.

Murhekar, M. V. and S. P. Zodpey. 2005. "Hepatitis B Virus Infection among Indian Tribes: Need for Vaccination Program," *Indian Journal of Gastroenterology*, 24: 269-270.

Oster, E. 2005. "Hepatitis B and the Case of the Missing Women," *Journal of Political Economy*, 113 (6): 1163-1216.

— 2006. "Explaining Asia's Missing Women: A New Look at the Data - Comment," *Population and Development Review*, 32(2): 323-27.

Oster, E., Chen, G., Yu, Xinsen and W. Lin. 2008. "Hepatitis B Does not Explain Male-Biased Sex Ratios in China," `http://home.uchicago.edu/ eoster/hbvnotecon.pdf`. Accessed February 09, 2009.

Sahni, M., Verma N., Narula D., Varghese, R. M., Sreenivas, V., and J. M. Puliyel. 2008. "Missing Girls in India: Infanticide, Feticide and Made-to-Order Pregnancies? Insights from Hospital-Based Sex-Ratio-at-Birth over the Last Century," *PLoS ONE*, 3(5).

Sen A. 1990. "More than 100 million Women are Missing." *New York Review of Books*, December 20.

— 1992. "Missing Women: Social Inequality Outweighs Women's Survival Advantage in Asia and North Africa," *British Medical Journal*, 304: 587-88.

— 2003. "Missing Women Revisited," *British Medical Journal*, 327: 1297-98.

Ulizzi, L. and L. A. Zonta. 1995. "Factors Affecting the Sex Ratio in Humans: Multivariate

Analysis of the Italian Population," *Human Biology*, 67:1.

Table 1: Probabilities of Male Birth: India, 1992

	Estimate	**Std Err**	**t-stat***
Unrestricted Model			
BIRTH PARITY 1	0.555	0.002	18.233
BIRTH PARITY 2	0.547	0.002	14.626
BIRTH PARITY 3	0.542	0.003	10.357
BIRTH PARITY 4	0.543	0.004	8.167
BIRTH PARITY 5	0.538	0.005	4.526
BIRTH PARITY 6	0.537	0.009	2.594
Restricted Model			
	0.547	0.001	25.247
N=44288			
Likelihood Ratio Test			
LR-Stat = 60.257			

*t-stat tests the equality of the estimate to 0.514

Table 2: Probabilities of Male Birth: India, 1998

	Estimate	**Std Err**	**t-stat***
Unrestricted Model			
BIRTH PARITY 1	0.547	0.002	15.4871
BIRTH PARITY 2	0.548	0.002	16.036
BIRTH PARITY 3	0.546	0.003	12.544
BIRTH PARITY 4	0.540	0.004	7.326
BIRTH PARITY 5	0.543	0.005	5.486
BIRTH PARITY 6	0.541	0.009	2.877
Restricted Model			
	0.546	0.001	25.730
N=49817			
Likelihood Ratio Test			
LR-Stat = 53.546			

*t-stat tests the equality of the estimate to 0.514

Table 3: Probabilities of Male Birth: India, 2005

	Estimate	**Std Err**	**t-stat***
Unrestricted Model			
BIRTH PARITY 1	0.537	0.002	12.081
BIRTH PARITY 2	0.539	0.002	12.665
BIRTH PARITY 3	0.544	0.002	11.902
BIRTH PARITY 4	0.538	0.004	6.859
BIRTH PARITY 5	0.531	0.006	3.164
BIRTH PARITY 6	0.552	0.010	3.886
Restricted Model			
	0.539	0.001	21.184
N=61841			
Likelihood Ratio Test			
LR-Stat = 38.457			

*t-stat tests the equality of the estimate to 0.514

Table 4: ML Estimation Results: India, 1992

	Estimate	Std Err	t-stat
Interaction			
CONSTANT	0.04	0.01	4.95
Son Targeting			
CONSTANT	−1.07	0.10	−10.74
AGE	−0.01	0.00	−3.66
EDUCATION	−0.01	0.00	−2.74
WORK	0.05	0.02	2.27
CONTRACEPTION	0.55	0.02	23.38
RURAL	0.06	0.03	2.12
JOINT FAMILY	0.03	0.03	0.79
HINDU	0.08	0.04	1.98
POOR	0.17	0.08	2.06
MIDDLE INCOME	0.02	0.08	0.24
Family Size			
AGE	0.02	0.00	67.24
EDUCATION	−0.02	0.00	−27.32
WORK	−0.02	0.01	−2.76
RURAL	0.04	0.01	6.05
POOR	0.57	0.01	44.72
MIDDLE INCOME	0.53	0.01	43.61
HINDU	0.15	0.01	15.97
Probability of Male Birth			
CONSTANT	0.34	0.03	10.74
AGE	−0.01	0.00	−11.22
EDUCATION	−0.01	0.00	−6.34
WORK	−0.02	0.01	−3.18
CONTRACEPTION	−0.07	0.01	−9.01
RURAL	0.01	0.01	1.02
JOINT FAMILY	0.03	0.01	2.38
HINDU	−0.03	0.01	−2.35
POOR	0.00	0.03	0.07
MIDDLE INCOME	−0.00	0.02	−0.14
TRIBAL	0.00	0.01	0.27
Loglikelihood = -184164.8			
N = 43282			

Table 5: ML Estimation Results: India, 1998

	Estimate	**Std Err**	**t-stat**
Interaction			
CONSTANT	0.05	0.01	5.90
Son Targeting			
CONSTANT	−1.10	0.10	−11.64
AGE	−0.01	0.00	−3.86
EDUCATION	0.00	0.00	0.63
WORK	0.12	0.02	5.63
CONTRACEPTION	0.59	0.02	25.27
RURAL	0.06	0.03	2.37
JOINT FAMILY	0.06	0.03	2.19
HINDU	−0.16	0.03	−6.19
POOR	0.20	0.08	2.58
MIDDLE INCOME	0.07	0.07	0.93
Family Size			
AGE	0.02	0.00	71.19
EDUCATION	−0.02	0.00	−32.13
WORK	0.00	0.01	0.05
RURAL	0.04	0.01	7.31
POOR	0.61	0.01	47.30
MIDDLE INCOME	0.54	0.01	45.16
HINDU	−0.06	0.01	−9.17
Probability of Male Birth			
CONSTANT	0.28	0.03	9.40
AGE	−0.00	0.00	−9.65
EDUCATION	−0.00	0.00	−4.64
WORK	−0.04	0.01	−5.44
CONTRACEPTION	−0.09	0.01	−11.68
RURAL	0.02	0.01	2.45
JOINT FAMILY	0.05	0.01	4.96
HINDU	0.02	0.01	3.03
POOR	0.01	0.02	0.46
MIDDLE INCOME	−0.00	0.02	−0.02
TRIBAL	−0.02	0.01	−2.29
Loglikelihood = -202268.2			
N = 49104			

Table 6: ML Estimation Results: India, 2005

	Estimate	**Std Err**	**t-stat**
Interaction			
CONSTANT	0.06	0.01	7.59
Son Targeting			
CONSTANT	−0.88	0.06	−14.29
AGE	−0.01	0.00	−5.18
EDUCATION	0.01	0.00	4.27
WORK	0.13	0.02	6.94
CONTRACEPTION	0.71	0.02	32.06
RURAL	0.00	0.02	0.20
JOINT FAMILY	0.01	0.03	0.32
HINDU	−0.15	0.02	−6.78
POOR	−0.02	0.04	−0.48
MIDDLE	−0.08	0.04	−2.43
Family Size			
AGE	0.02	0.00	100.52
EDUCATION	−0.02	0.00	−30.68
WORK	0.01	0.01	1.77
RURAL	0.06	0.01	12.38
POOR	0.45	0.01	45.01
MIDDLE	0.35	0.01	39.93
HINDU	−0.07	0.01	−12.89
Probability of Male Birth			
CONSTANT	0.21	0.02	10.03
AGE	−0.00	0.00	−6.53
EDUCATION	−0.00	0.00	−0.93
WORK	−0.04	0.01	−7.05
CONTRACEPTION	−0.09	0.01	−11.66
RURAL	0.01	0.01	1.62
JOINT FAMILY	0.03	0.01	3.48
HINDU	0.02	0.01	2.97
POOR	−0.00	0.01	−0.30
MIDDLE	−0.01	0.01	−0.72
TRIBAL	0.01	0.01	1.17
Loglikelihood = −237659.8			
N = 60071			

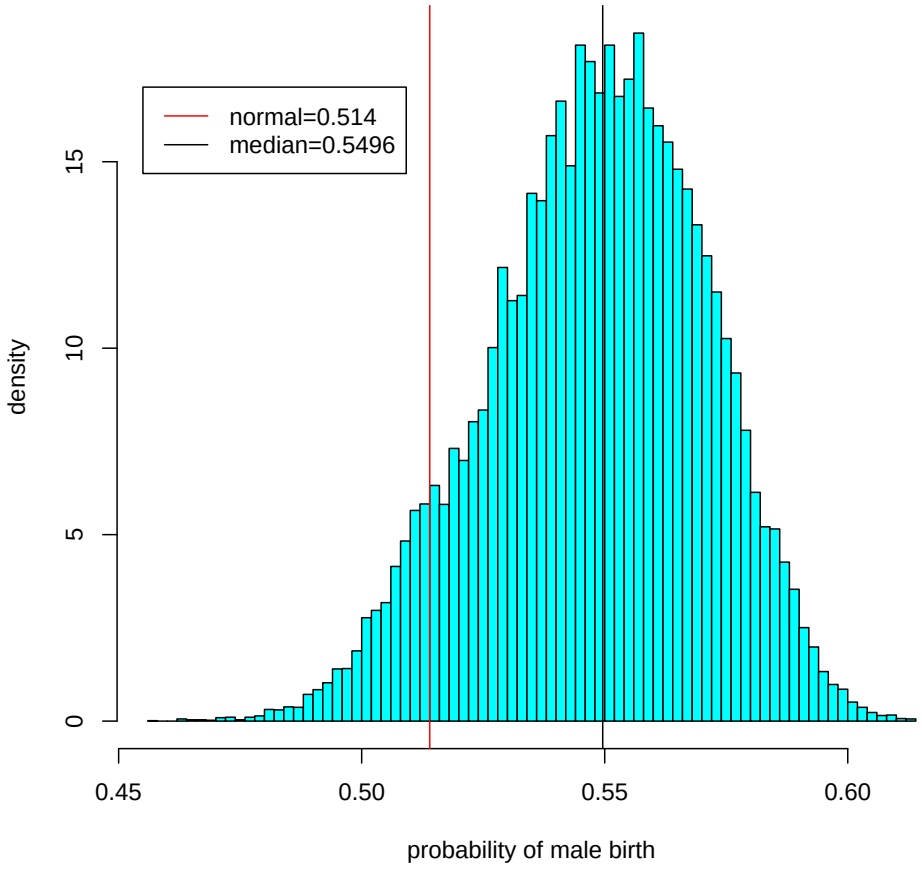

Figure 2: Predicted Probability of Male Birth, 1992: Full MLE

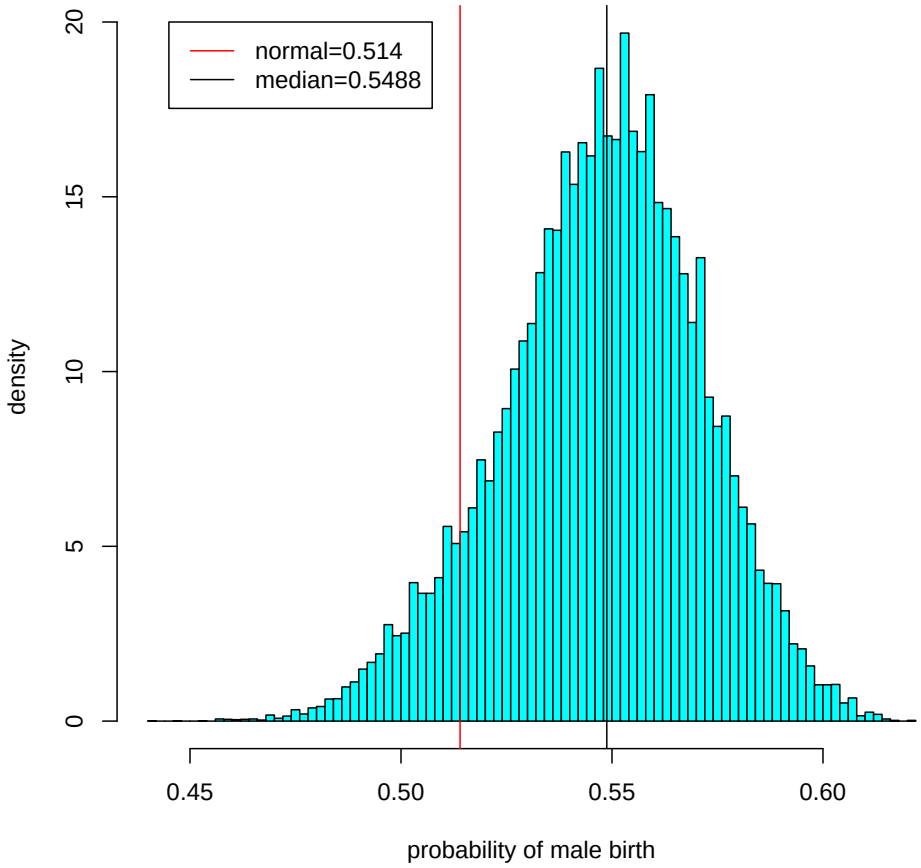

Figure 3: Predicted Probability of Male Birth, 1998: Full MLE

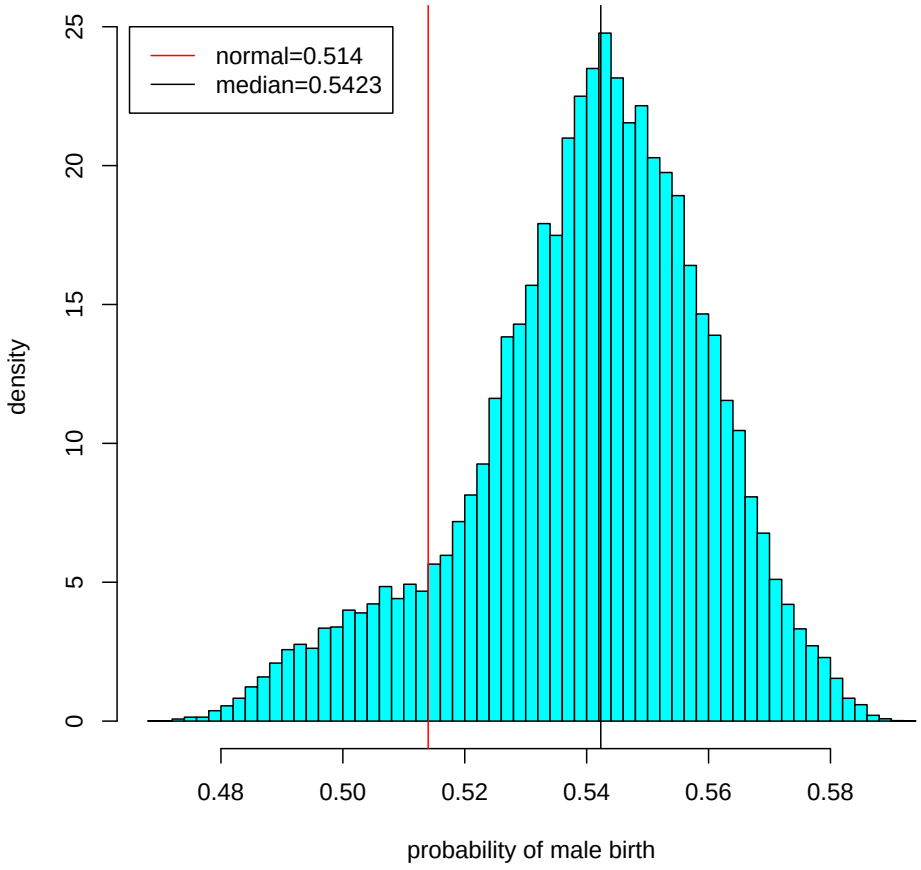

Figure 4: Predicted Probability of Male Birth, 2005: Full MLE