Son Preference, Differential Stopping Behavior, and the Fertility Outcomes

Sindhu Vasireddy

Updated: September 7th, 2018

University of Hawai'i-Manoa

East-West Center

Son Preference, Differential Stopping Behavior, and the Fertility Outcomes

Sindhu Vasireddy*

This paper studies son preference through gender biased differential stopping behavior, and the impact of it on fertility outcomes in India. Existing literature indicates that parents with a high preference for sons use contraceptive methods to halt fertility following a male birth. I find that this differential stopping behavior favoring male children leads to reproductive decision-making where people decide to have a higher number of pregnancies, births, and lower contraceptive usage if the first-born child is female, leading towards girls being concentrated in larger families when compared to boys. These fertility outcomes prevail across households with lower fertility rates and differ by the age of the mother. I find further evidence in terms of health outcomes to show that having an older brother improves survival chances of the second children, especially if the second child is female.

*Ph.D. Program in Economics, University of Hawai'i-Manoa (USA) and affiliated with the East-West Center. Address: 1711 East-West Road, 1206I Hale Manoa, Honolulu, Hawaii, USA -96848-0001. Email: vs47@hawaii.edu. Acknowledgments: For developing the ideas presented here, I have received helpful inputs from Timothy Halliday, Andrew Mason, Sang-Hyop Lee, Kyeongkuk Kim, Roman Shishkin, Abhishekaran Karapurath, and seminar participants of the Applied Microeconomics group, University of Hawai'i-Manoa.

1. Introduction

This paper tries to understand how fertility outcomes are impacted by son targeting differential stopping behavior (henceforth DSB). There is a rich literature on differential stopping behaviour and sibling effects in India, but the closest one that comes to the estimates of this analysis is the study by Basu and De Jong (2010) in India, who quantify the sibling effects across Asia, South Asia, North-Africa and Sub-Saharan Africa based on the fertility history of mothers in 1992. Their results indicate that girls have a larger number of siblings across all the mentioned countries and are born at relatively earlier birth parities.

The results of this study concur with their estimates and provide evidence for son-targeting fertility behavior using the sibling effect too. However, this paper differs from Basu and Jong(2010) in three different ways. Firstly, I extend the study to understand sibling effects across mothers of different ages and differentiate the impact of son targeting behavior by fertility outcomes such as the number of pregnancies, births, contraceptive usage and pregnancy termination. For this, I use the the identification mechanism followed by Altindag(2016) to estimate the sex ratios in Turkey and Jensen(2003) in India, by using the gender of the first child as a random shock and observe gender differentials by the variation in the couples' response to it. There is a vast amount of literature[1] indicating that India uses ultrasound scanning and abortive technology to terminate pregnancies if the fetus is female, which also necessitates that we identify the fertility outcomes separately to understand how each variable responds to gender biased fertility decision-making. For instance, we observe a greater increase in the number of pregnancies when compared to the growth in the number of living children, if the first child is female.

[1] For recent discussions, refer to Bhalotra et. al. (2010), Anukriti et. al. (2016)

Secondly, I show evidence for the differential contraceptive usage in the households following a female birth. Understanding contraceptive usage based on son preferences helps us in identifying the heterogeneity of fertility responses based on the household and individual characteristics of the mother. This also helps in accounting for fertility choices when the actual fertility behavior differs from the desired proportions that Basu and Jong (2010) use, especially since the desired outcomes are heavily understated in the National Family and Health Survey data when compared to the actual fertility behavior (Jayachandran, 2017).

And lastly, I expand this work to study the fertility behavior of households from 1992 to 2006 and find a consistent presence of sibling effect across fertility groups of different sizes based on gender differences. As the Indian government has been trying to encourage more people towards smaller families through family planning centers, increasing the knowledge and access to birth control, studying how the fertility outcomes are impacted by gender preferences in the presence of groups showing lower fertility levels gains importance.

There has long been a debate on whether there is an impact of Son Preferred Differential Stopping Behavior (hereby SP-DSP) on sex ratios. A recent study by Perwez et al. (2012) showed that the stopping rule behavior plays a major role in skewed sex ratios in India using empirical evidence, contradicting which Dreze (2012) showed that their arguments are without a basis and termed them 'old fallacies in new bottles' using the law of large numbers in a large population. Srivastava (2012) concurred with Dreze (2012) and further demonstrated that if each birth is an independent event with a constant probability of having a girl, the decisions of families on when to stop reproducing have no effect on the population sex ratios. My model agrees with the same and adds to it by showing how differential stopping behavior may result in uneven distribution of male and female children across heterogeneous households.

Literature also supports the idea that this sibling effect impacts children through allocative behavior. Notably, Jensen (2003) finds that son targeting behavior leads to higher fertility rates overall, resulting in the households having to share a limited amount of resources over a large family size. He finds that this behavior results in about a one-tenth to one-quarter of differences in educational outcomes based on the size of the siblings.

Park and Cho (1983) show that the sex ratios of siblings in small families are skewed in favor of boys, and that sex at the last birth is highly correlated with the decision of having an additional child in Korea. In this paper, I find reasonable evidence that supports a similar idea that the probability of the last child being male reduces the desire for more children through contraceptive usage (Appendix Table A2).

Other studies by Dahl and Moretti (2008) show that the number of children within a family is significantly higher in households with a first-born daughter when compared to those with firstborn sons in the U.S.. Their studies prove that son-preferring behavior is prevalent throughout the world in myriad formats, which Dahl and Moretti (2008) investigate by stringing various pieces of evidence together.

The findings in this paper contribute to the literature by providing evidence for son-targeting fertility behavior and find that there is a larger likelihood of girls to be concentrated in large families, while compared to the boys and add across different fertility groups and age groups of mothers. This goes to show that gender inequalities resulting from the sibling effect have prevailed across time. I further explain the health outcomes of such a composition and find that having an older brother improves survival chances of the second children, especially if the second child is female.

This goes to show that there is a need to provide incentives for female births, especially when there is a girl following another girl. More schemes like Ladli Beti in Haryana that provide incentives for having two girls need to be encouraged. Further, few schemes in India are open to households with higher education and income levels which encourage the birth and survival of girls. Our results from Table 4, when added to the results of the Appendix table A.1 clearly show a higher preference for boys when compared to girls among the wealthy sections and those within low fertility quintiles, where most of the educated women fall, indicating a need to address the gender preference and sibling effects in the upper classes either through more incentivizing programs or through higher female empowerment programs.

The rest of the paper is organized as follows. Section 2 describes the Indian context with reasons and mechanisms leading to gender biased outcomes. Sections 3 and 4 discuss data and the empirical strategy. Section 5 presents results, and Section 6 presents estimates of the sibling effect on a few health outcomes for children who are born in the second parity. Section 7 concludes.

2. Context

The effect gender bias on women has first been brought to public attention by Sen (1990) in his essay indicating that more than a hundred million women are "missing" through sex-selective abortion and excess female mortality. Son preference has been prevalent for centuries in India, and this bias is attributed to a number of reasons like the continuing social practice of dowry system, where the bride's family pays a mutually bargained amount to the groom's family for marrying their daughter, along with bearing the wedding expenses. As a result of this, daughters are often associated with high economic costs for the parents (Das Gupta, 2003).

Other reasons for parents associating daughters with high economic costs could be due to low participation of women in the workforce in India. A recent estimate by the World Bank Report on Gender Inequality (2011) that measures the participation of women in the workforce, ranks India 135^{th} out of 144 countries of the world, lower than many countries in Sub-Saharan Africa and the Middle East. Women on the average sum to less than 20 percent of the total workforce in the nonagriculture sector and 27 percent in total (Indian Statistical Department).

Apart from economic reasons are the religious beliefs of Hindus: Moksha and Kanya Daan. While the former prescribes that souls of the deceased reach heaven only if a son or a grandson lights a funeral pyre, the latter is a customary practice of sending daughters to live with the in-laws after marriage, while the son is expected to live with his parents. Since it was predominantly the son's responsibility to look after the parents, he was also the only one entitled to family wealth according to the Hindu customs. It was only after the Hindu Succession (Amendment) Act in 2005 that Hindu daughters gained the right to seek an equal share in ancestral property. Apart from the reasons mentioned above, numerous cultural reasons add up to hold the patriarchal structure of Indian society, where a higher percentage of people see incentives in having a male child when compared to having a female child. Around the late 80s mobile ultrasound units started touring rural Haryana, with posters advertising "Spend 500 rupees now and save 50,000 later", referring to the cost of pre-natal sex-determination as compared to the expense of a dowry.

With the accessibility of ultrasound scanning, more people were able to determine the gender of the fetus leading to sex selective abortions in India. (S. Bhalotra et al., 2010). Other forms of gender bias came through differential treatment in ante-natal care, lesser likelihood of receiving immunization and nutrition, or girls being less likely to be taken to the hospital during an illness reducing the childhood survival rates of girls from birth to infancy. (S. Jayachandran, 2017; Choi and Lee, 2006; Lleras Muney, 2010) In spite of girls being endowed with higher disease resistance rates that should improve their chances of survival[2], the male to female sex ratios are high in India. (Figure 1).

Improved education of mothers, and the government efforts to reduce population growth by the extensive establishment of family planning centers, campaigns suggesting that a 'small family is a happy family' airing over the televisions and posters for decades, increased maternal education resulting in a quality-quantity trade-off, changing preferences among other things, have led to an increase in the use of contraception over the years. Couples in India are increasingly using contraceptive methods, which has risen from 13% in 1993 to 48% of the households by 2006[3].

These changing preferences have led to an increase in the male-to-female sex ratio as Jayachandran (2017) shows in her work[3]. She estimates that fertility decline can explain roughly one-third to one-half of the sex ratio increase in Haryana. This paper contributes to this stream of literature by quantifying the asymmetry in how much households ensure the survival of more boys versus more girls, based on the gender of the first child across fertility quintiles.

3. Data and Descriptive Analysis

There are three important criteria that the data needs to meet to conduct our study. Firstly, the data must contain the complete and detailed fertility histories of female respondents, along with those living in the household as well as those who separated from the families. Secondly, we need to understand the gender preferences within the households. Lastly, we need information on

[2] There is a large amount of medical literature on this topic. To name a few, refer to Lozano (2010), Anker (2007) [3] Estimates calculated from National Family and Health Survey data used in this study.
[3] Suggests that households use stopping behavior for smaller families and the sex selective abortion technology for sex ratio widening.

household and individual characteristics that will help us in controlling for and analyzing the heterogeneity of responses across socio-economic groups.

Since NFHS surveys have been conducted with the primary aim of providing information on maternal, child and reproductive health, they contain data on the entire fertility history of families, information on contraceptive usage and awareness among ever-married women aged 19-49 that are of interest to us. This included information on the number of sons and daughters who were living in the household had left the household or had died; the birth order of each child; if they were using or not using a contraceptive and the reasons for it.

NFHS-1 contains information from surveys of 88,562 households of ever-married women between the ages of 13 and 49, NFHS-2 includes information from a population of 91,196 households of ever-married women age 15-49 and NFHS-3, survey data of 109,041 households. The large size and detail of this data enable us to meet the sample size requirements better than some previous studies. The surveyors also collected information on various demographic and socioeconomic characteristics that can affect a woman's reasons for discontinuing contraception which is studied to understand heterogeneity in son preference and contraceptive usage. These variables are residence (urban, rural); education maternal and paternal (in highest years of education); religion (Hindu, Muslim, others); and geographic region (South and Central India). These predictors are known to have substantial effects on contraceptive use (Ramesh et al. 1996) and are likely to affect discontinuation and intention to not use contraception as well.

The sample size of this data varies from state to state. In some states the sample design was self-weighting, and in others, certain categories of respondents were over-sampled for various reasons. Analysis of data from these surveys required the usage of frequency weights to restore the correct proportions for a better representation of national level data.

I pool the three rounds of this NFHS conducted in 1992-93, 1998-99 and 2006-06 and include year dummies in the analysis. This dataset also addresses recall bias to a large extent by the usage of calendar tools that encourage respondents to link events chronologically and by a probing interviewer guiding to reveal gaps to help in addressing the inconsistencies in reported reproductive histories. In this data, I drop births higher than the order of six, and the estimation sample has about 190,000 mothers, with about 510,000 births remaining.

Table 1 presents the descriptive characteristics of the variables used in this estimation. Since there is a considerable amount of difference between the characteristics of the urban and rural areas of India, they are reported separately in the table. In rural areas, the sample means indicate a higher percentage of mothers with little or no education, of Hindu religion, 35 percent of them belonging to the scheduled castes (SCs) and Scheduled Tribes (STs) and get married at younger ages. The households in rural areas also lack access to provisions such as electricity and water. But they seem to have a higher percentage of girls when compared to the urban areas.

From Figure 1, we can observe that our data is representative of the census estimates. The overall sex ratios of both the Census and NFHS estimates are higher than the Child sex ratios, indicative of higher survival chance for girls once they cross their childhood.

4. Methodology

I try to identify the causal effects of son preference on fertility decisions by assuming that the gender of the first child is a random draw. We study the differentials in the reproductive decisionmaking of the couples as a response to this exogenous shock.

A reduced form of the baseline equation is:

$$F_{mtr} = \alpha + \beta X'_{i} + \tau Z_{irt} + \gamma_r + \delta_t + \omega_{tr} + {}_{mtr} \quad (1)$$

Where, F_{mtr} stands for the fertility outcomes (being the number of pregnancies, births, living children, contraceptive usage, and pregnancy termination) for mother *m*, interviewed at time period *t*, living in the region *r*. *X* is a vector of family characteristics (age of the mother, age at first marriage, maternal education in years, paternal education in years, if they are residing in a rural area). *Z* stands for the gender of the first child. It takes the value of 1 if the first child is a female, and the value of a 0 if the first child is male. γ_r controls for region fixed effects and δ_t controls for the year of the survey. ω_{tr} refers to the region-specific year effects[4].

I use OLS and Poisson Likelihood functions for estimating equation (1). OLS is used for baseline estimation, and Poisson estimates report the percentage changes in sibship size induced by a female birth. Poisson estimation was specifically chosen since it an efficient tool for measuring the count values of the number of children, and since it secures positive outcomes.

There are two major exclusion criteria that I follow in this analysis. Firstly, Z is restricted to surviving, singleton births to isolate gender responses from responses to the death of the first child, or twinning outcomes since the Poisson model does not support group outcomes. Secondly, I restrict the responses to 86 % of births that occurred in mothers within current locations to address the problem of migration.

[4] Fertility decisions are strongly influence by social norms and diffusion effects. Capturing social effects in householdlevel analyses raises serious conceptual and empirical difficulties, to control for which, we use region and time fixed effects.

One major concern in this paper is the prevalent usage of ultrasound scanning in India for prenatal sex-selective abortions. Sex-selective abortions in the first birth will lead to the failure of the assumption of exogeneity of the key instrument used in this analysis.

As Altindag (2016) suggests, we can compare the family characteristics of those who have first-born sons, with those who have first-born daughters, and further use the regression below to run a logit model test and check the joint χ^2-test for the null hypothesis that all the estimated coefficients in the right-hand side of the equation are jointly equal to zero.

$$Z_{irt} = \lambda + X'_i\phi + \gamma_r + \delta_t + _{mtr} \quad (2)$$

From further robustness checks and specification tests, we find that these may be able to explain more than 60 percent of the sibship size (***Table 2***).

5. Results

(Table 3)

Results of the pooled OLS regressions and the Poisson distribution MLE are shown in Table 3. Estimates from every panel are separate regressions run by the age of the mother since this is one of the most important determinants of fertility[5] and because fertility outcomes greatly vary by age.

Pooled OLS estimate in the first column indicates that having a first-born female child have 0.22 more pregnancies for mothers within the age group of 15-49. The MLE estimate in the first column shows that families with daughters born in the first birth are likely to have nearly 7.1 per cent of children more than those with sons born in the first parity. OLS and MLE estimates from column (3) panel (B) show that mothers between the ages of 15-29 have 0.74, or 3.8 percent more

[5] Determinants of fertility in the unadjusted outcomes have been derived from Visaria (2004)

children if they have a daughter in the first birth. These estimates increase gradually by age showing a 0.245 and 0.240 more children referring to about 8.7 percent and a 5.8 percent increase in the sibship size.

We can also observe that this having a girl in the youngest age group leads to a negative usage of contraception[6] and a small and insignificant outcome of pregnancy termination. Mothers in this age group also have the least number of births induced by first-born daughters. Mothers in the youngest cohort may have not completed their fertility cycle, and these outcomes could be a result of this behavior.

From column (5) it can be noticed that if the first-born is a girl, abortion is not a statistically significant mechanism in the younger cohort. However, this outcome increases slightly over time. An important point in understanding this column is that this pregnancy termination could have resulted out of a miscarriage or other health concerns for the mother and child and need not necessarily refer to sex-selective abortion. This outcome can also be a result of underreporting behavior since ultrasound scanning that leads to sex-selective behavior has legally been banned in India under the Prenatal Sex Diagnostic Techniques (Regulation and Prevention of Misuse) Act (PNDT) in 1994. This act has been advertised and came into full force in 1996.[7]

5.1. Heterogeneity

(Table 4)

[6] Contraceptive usage includes modern techniques being, female sterilization, male sterilization female condoms, male condoms, pills, foam, vaginal douche, intrauterine devices and so on., and traditional techniques such as withdrawal, abstinence.

[7] More details on PNDT can be found in Visaria (2005), Retherford and Roy(2003)

A major amount of literature on son preference and differential stopping behavior in India focuses on understanding heterogeneity by constructing an index of baseline characteristics that predicts the values of outcomes in the absence of treatment. However, studies by Abadie et al. (2013) show substantial biased estimates of treatment effects that result from using this technique and suggest using the 'leave-one-out'(hereby LOO) and 'repeated split sample'(hereby RSS) estimators to correct the bias. While RSS divides the sample into two groups and solves the problem of overfitting by repeating the prediction of one sample over several iterations, the LOO method solves the problem of over-fitting by excluding each observation when estimating the coefficients used to calculate its own predicted value[8].

Sibling Effect as defined by Basu and Jong (2010) refers to the difference between the expected increase in the number of children if the first child is a girl when compared to an increase in the expected number of children if the first child if a boy. To estimate this, the RSS and LOO methods begin by using a set of covariates to predict fertility outcomes for the untreated group. These estimates are then used to predict results for the full sample. Then the outcomes are stratified into quintiles, where the treatment effects are estimated. Instead of interacting each regression outcome with a treatment dummy, this method of estimation gives an index of predicted outcomes by using all the relevant covariates[9].

Robustness checks have been performed in this paper by dropping covariates in the unadjusted panels of the RSS and adding them back in the adjusted sections**.** Table 4 presents us with the

[8] For further information, refer to Abadie et al. (2013)

[9] Adjusted regressions include covariates for the mothers age, mothers age at marriage, maternal education, paternal education, rural residence, Hindu and Muslim dummies, region dummies, year dummies, region-specific yeareffects, wealth indicators (water connection to the household, electricity connection, good housing and a car), decision-maker for contraceptive usage, and an indicator for being in under privileged caste groups (includes Scheduled Castes and Scheduled Tribes).

results of this estimation. This shows the effects of a first-born daughter on the number of siblings at each fertility quintile. While the unadjusted predictions in columns (1) and (3) use the determinants for fertility being maternal and paternal education, contraceptive usage, mothers age at marriage, and rural residence; the adjusted regressions in columns (2) and (4) implement a full set of controls including year dummies, mother's age, mother's age at first birth, religion, scheduled caste, scheduled tribe, husbands age, contraceptive decision-making power and wealth indicators based on water and electricity connection to the household, solid housing and the possession of a car. There is no difference in the direction, and little difference in the magnitude between adjusted and unadjusted outcomes indicating that irrespective of the determinants used, the gender of the first child is exogenous.

Column (5) presents the mean number of children by fertility quintile if the first child is a boy. We can observe from this column that the mean number of children vary by fertility outcomes from about 2 children in the lowest fertility quintile to 4 children in the highest. It can be further observed from column (4) and (6) that there is a high and statistically significant increase in the number of children if the first child is female at different fertility outcomes. If the first child is a girl, then the number of children increase by 0.223 children, in the third quintile, which amounts to 8.8 percent increase in the number of children if the first child is female.

(Figure 2)

As the fertility levels increase in our estimation, the expected number of siblings for girls who are first children spikes up when compared to the expected number of siblings if the first child is a boy till the third quintile. However, from the third quintile, though there is a substantial increase in the mean number of children, there is little increase in the expected number of children if the first child is female, leading to a lower sibling effect.

5.2. Specification tests:

5.3. Randomness of the sex of first births:

A key assumption we use as our identification strategy is the randomness of gender of the first child. If our assumption fails, the causal effect of sibship composition on contraceptive usage becomes questionable. (A) There is a proof for the validity of this test in ***(Table 4)***, as explained in the section above. (B) It can be can see from ***(Figure 3)*** that the male-to-female sex ratios in the sample for the first order birth are within the "normal" range. Any increase in the number of male children is due to male children having a higher proclivity for conception. (C) Summary statistics from ***(Table 2)***, indicate very little differences in the family characteristics across genders at the first parity. (D) Field studies in Gujarat and Haryana by Visaria (2004) showed that the sex ratio of children born in the first parity was within the normal range even after the prevalence of ultrasound scanning tests. (E) Bhalotra et. al. (2010) show further proofs for the validity of this assumption for first order births in the presence of ultrasound scanning and prenatal sex-selective technology.

5.4. Health Outcomes:

I further try to understand how the previous sibling's sex changes the gender health gap in the next parity, for which I use the differences-in-differences approach, with an exogenous variation being the first child's gender. The reduced form of the equation is:

$$H_i = \beta_0 + \beta_1 Z_{i1} + \beta_2 Z_{i2} + \beta_3 (Z_{i1} \times Z_{i2}) + \eta_i$$

Where H stands for the health outcomes being Infant Mortality and Child Mortality; Z_{i1}indicates the gender of the second child whose mortality outcome is in question, taking a value of 1 if the child is female, and 0 is the child is male. Z_{i2} indicates the gender of the first child which is 1 if the child has an older sister and 0 if the child has an older brother.

Here, $(Z_{i1} \times Z_{i2})$ takes a value of 1 if both the first and second children are female, and β_3 is positive if having a girl in the first-birth causes the boy to be more valuable in the second birth. β_1, indicates the sibling effect in infant mortality outcomes if the first child is male, and $(\beta_2 + \beta_3)$ shows the sibling effect if the first-born is female.

Differential Stopping Behavior can not only cause differences in sibship composition, but may also increase a competition for limited resource unevenly among the genders. As Jensen (2003) shows in his work that this composition leads to disparities in educational attainment. I show that the health outcomes also differ by the gender of the first child.

(Table 5.A, Table 5.B)

Most of the couples in India have at least 2 children, and in Table 5, I observe the mortality rates of the second sibling based on the gender of the first child. Table 5.A. studies the infant mortality rates and it can be observed that there is a significant different between the children with an older sister when compared to older brothers. Girls with an older sister are 2.6 percentage points less likely to survive before reaching the age of 1 when compared to boys with an older sister. Though girls with older male siblings experience higher infant mortality than boys with older brothers, this difference is not statistically different.

In both the cases, whether the second is a girl or a boy, boys enjoy lower mortality rates than girls. However, existing studies show that girls have a higher amount of disease resistance and immunity after birth in the absence of differential treatment if they are born in homogenous families. So these mortality differences must either be a cause of either second born girls being born in different families with a lesser potential to care for them, or the girls are experiencing differential treatment or a bias in the allocation of resources.

The results in table 5.B showing the child mortality outcomes for second-born children corroborate this evidence. Mortality of the second child is least when there are children of either sex present in the two children, that is, in the GB and BG combinations. However, mortality increased when there are two boys and two girls alive, such as BB and GG. Having an older brother means that either of the children have nearly an equal likelihood for survival, while we huge disparities in the childhood mortality between a second child who is a girl and a second child who is a boy if they have an elder sister. When there is an elder sister, girls are 1.4 percentage points less likely to survive than boys.

Having an older sister increases infant mortality by about 2 percentage points and childhood mortality by about 1.7 percentage points. These differences are significant at both the adjusted and unadjusted regressions.[10]

[10] Adjusted regressions include covariates for the mothers age, mothers age at marriage, maternal education, paternal education, rural residence, Hindu and Muslim dummies, region dummies, year dummies, region-specific yeareffects, decision-maker for contraceptive usage, and an indicator for lower caste (includes Scheduled Castes and Scheduled Tribes).

6. Conclusion

Son Preferring differential stopping behavior has remained persistent through time and despite declining fertility levels across all fertility quintiles. This can be seen from our results indicating lower contraceptive usage after a first-born daughter when compared to a son. Son preference thus leads to reproductive decision-making where people decide to have a higher number of pregnancies, births and living children if the first-born child is female. This behavior results in girls being concentrated in larger families when compared to boys.

This can have adverse outcomes for females since a higher number of siblings for girl can lead to an increased competition for resources irrespective of the presence of differential treatment between girls and boys in resource allocation in the household. I also find that these fertility outcomes by the age of the mother ranging from a 3.8 percent increase in the family size if the first born is a girl in younger cohorts to nearly a 9 percent increase in the family size among middle-aged cohorts. The gender bias outcomes when further grouped into fertility quintiles increase over the till the mid-range and decrease in the later stages indicating a lower discriminatory behavior among women who choose to have large families.

Further studies indicate that having an older brother improves the survival chances for girls to a large extent while being a boy in the second birth, irrespective of the gender of the sibling results in lower mortality outcomes. These results are especially disturbing since existing medical literature clearly points towards a higher chance of being susceptible and suffering a greater severity of illness due to infectious diseases during both infancy and childhood for male children when compared to females.

There is a strong need for policy intervention that aims at reducing son targeting fertility behavior and can encourage the placement of girls in smaller families, where there is a lesser competition for resources. Additionally, we need more incentive programs to encourage the birth and survival of a girl child, following another girl.

Also, most of the current policy interventions aimed at improving the survival and welfare of girls in India are conditionally open only to those belonging to the lower income and underprivileged castes. Our results indicate that son preference, differential stopping behavior, and adverse sibling effects are present among the lower fertility quintiles that are characterized by high maternal education and wealth throughout literature, indicating a need to intervene and encourage the birth and survival chances of girls in the upper classes either through open incentivizing programs or through female empowerment.

References

Abadie, A., M. M. Chingos, and M. R. West (2013). "Endogenous stratification in randomized experiments," NBER Working Paper No. 19742.

Altindag, O., (2016). "Son Preference, Fertility Decline and Non-Missing Girls of Turkey," Demography. Apr: 53(2):541-66.

Anker, M., (2007). "Addressing sex and gender in epidemic-prone infectious diseases," Geneva, Switzerland: World Health Organization.

Anukriti, S., S. Bhalotra, and H. Tam (2016). "On the Quantity and Quality of Girls: New Evidence on Abortion, Fertility, and Parental Investments," IZA Discussion Paper 10271.

Basu, D. and R. De Jong (2010). "Son targeting fertility behavior: some consequences and determinants," Demography, 47, 521–536.

Becker, G. S. and H. G. Lewis (1974). "Interaction between quantity and quality of children," Economics of the family: marriage, children, and human capital, UMI, 81–90.

Ben-Porath, Y., & Welch, F. (1976). "Do Sex Preferences Really Matter?" The Quarterly Journal of Economics, 90(2), 285–307.

Bhalotra, S. and T. Cochrane (2010). "Where Have All the Young Girls Gone? Identification of Sex Selection in India," IZA Discussion Paper No. 5381.

Cutler, D., Deaton, A., & Lleras-Muney, A. (2006). "The Determinants of Mortality," Journal of Economic Perspectives, 20(3), 97–120.

Dahl, Gordon B. and Enrico Moretti. (2008). "The Demand for Sons," Review of Economic Studies (75): 1085–1120.

Das Gupta, Mukherjee. (2003). :Arranged marriage, dowry and female literacy in a transitional society," CREDIT Research Paper No. 03.12.

Drèze, J. (2012). "Child Sex Ratio and Sex Selection Old Fallacies in New Bottles," Economic & Political Weekly: XLVII (38), 97–99.

Jayachandran, S. (2014). "Fertility Decline and Missing Women," American Economic Journal: Applied Economics, 9(1), January 2017, pp. 118-139.

Jensen, R. T. (2003). "Equal treatment, unequal outcomes? Generating sex inequality through fertility behavior," Mimeo, Harvard University.

Lozano R, Naghavi M, Foreman K, et al. "Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010," Lancet. 2012; 380:2095–128.

Mitra, A. (2014). "Son Preference in India: Implications for Gender Development," Journal of Economic Issues, 48(4), 1021–1037.

Park, C. B., and N.H. Cho (1995). "Consequences of son preference in a low-fertility society: imbalance of the sex ratio at birth in Korea," Population and Development Review, 59–84.

Rodriguez, G. (2007). "Poisson Models for Count Data. Chapter 4," Bernoulli, 1–50.

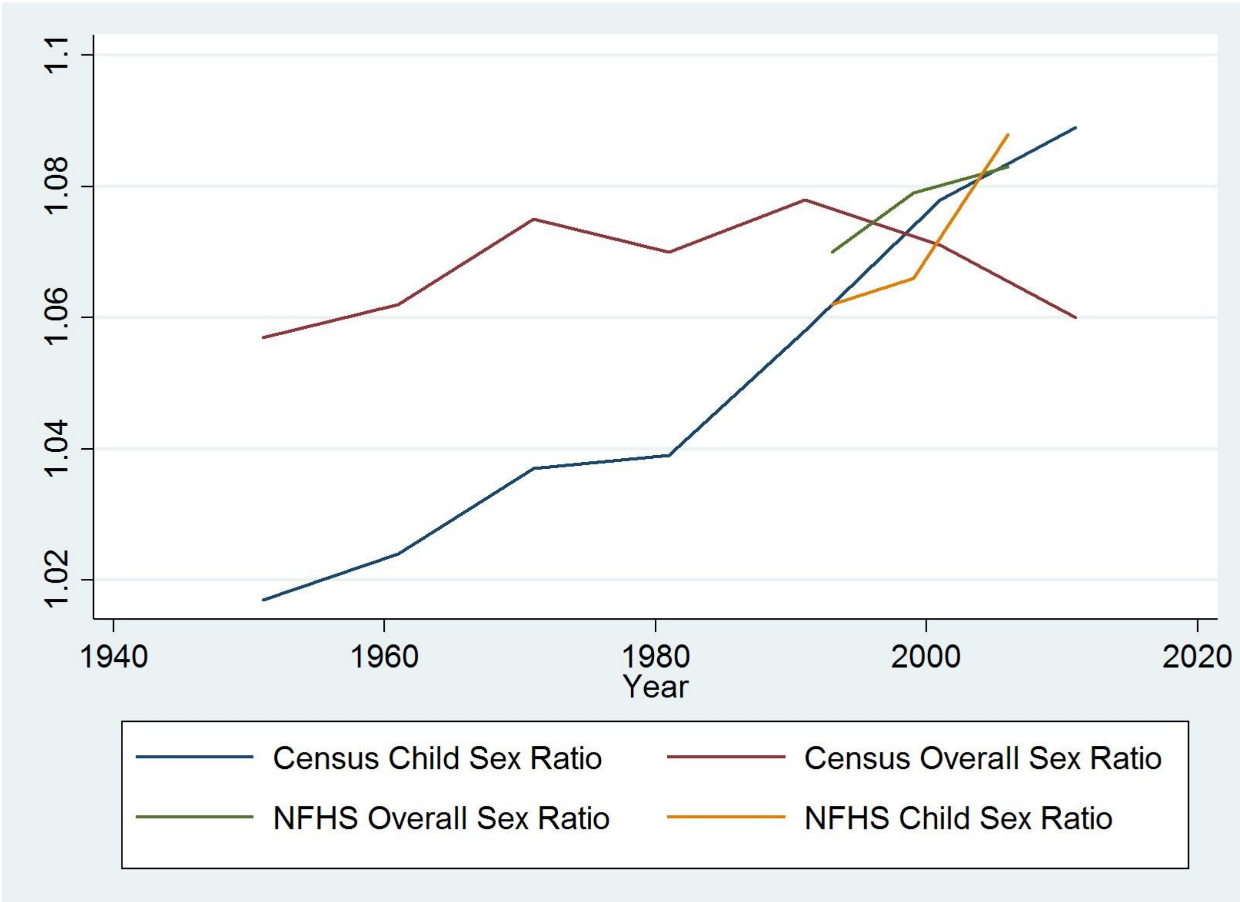
Sen, A. (1990): "More than 100 million women are missing," The New York Review of Books: Vol. 37, No. 20: 61- 66.

Shahid Perwez, Roger Jeffery and Patricia Jeffery (2012). "Declining Child Sex Ratio, Female Infanticide and Sex Selective Abortion in India: A Demographic Epiphany?", Economic and Political Weekly, 47 (33): 73-77.

Shrivastava, A. (2012). "Child Sex Ratio and Sex Selection," Economic & Political Weekly: XLVII (38), 97–99.

World Bank (2011). "World development report 2012: gender equality and development," World Bank Publications.

Figure 1: Childhood and over-all sex ratio trends



Note: Figure shows the estimated sex ratios under age five and overall sex ratio from the Census estimates of India from 1951 to 2011, and National Family and Health Survey Estimates for the years 1991-92, 1998-99 and 2005-06. X-axis is scaled to the years of data. Y-axis plots the sex ratio.

Table 1: Descriptive Statistics

	All India	Urban	Rural
Age	**29.14**	**31.53**	**30.27**
	(9.45)	(8.01)	(8.34)
Age at first Marriage	**18.98**	**19.63**	**18.24**
	(3.93)	(3.98)	(3.58)
Education			
None	**0.379**	**0.251**	**0.507**
	(0.46)	(0.433)	(0.499)
Finished Primary	**0.152**	**0.134**	**0.173**
	(0.349)	(0.340)	(0.378)
Finished Secondary	**0.373**	**0.458**	**0.291**
	(0.496)	(0.498)	(0.454)
Higher Education	**0.096**	**0.156**	**0.027**
	(0.308)	(0.363)	(0.162
Religion			
Hindu	**0.748**	**0.721**	**0.76**
	(0.444)	(0.448)	(0.426)
Muslim	**0.131**	**0.156**	**0.104**
	(0.337)	(0.362)	(0.306)
Christian	**0.076**	**0.075**	**0.078**
	(0.282)	(0.264)	(0.269)
Caste			
Scheduled Caste	**0.174**	**0.165**	**0.185**
	(0.377)	(0.371)	(0.388)
Scheduled Tribe	**0.122**	**0.075**	**0.17**
	(0.343)	(0.264)	(0.376)
Children			
Number of Children	**2.68**	**2.49**	**2.87**
	(1.36)	(1.27)	(1.41)
Daughters as a part of children	**0.447**	**0.416**	**0.478**
	(33.89)	(33.10)	(34.16)
Has Electricity	**0.758**	**0.915**	**0.604**
	(0.427)	(0.277)	(0.481)
Has Water	**0.413**	**0.444**	**0.374**
	(0.144)	(0.252)	(0.233)
Has Television	**0.146**	**0.218**	**0.124**
	(0.175)	(0.348)	(0.427)

Contraceptive decision-maker is Female	**0.114**	**0.103**	**0.124**
	(0.329)	(0.328)	(0.330)

Note: This table reports the baseline characteristics of the sample. The first column presents the covariate means for the households in India. The second and third columns present the covariate means differentiated by urban and rural characteristics.

Table 2: Baseline Characteristics of Families by First Child's Sex

First Child	*Boy*	*Girl*	*Difference*	*P-Value*	*N*
Mother's Age	31.01	31.49	0.48	0.51	30628
Mother's age at first marriage	18.66	18.49	-0.17	0.19	30628
Maternal Education	3.59	3.755	0.165	0.16	30628
Central	0.185	0.177	-0.008	0.66	30628
South	0.232	0.231	-0.001	0.59	30628
Rural	0.499	0.502	0.003	0.48	30628
Father's Age	36.27	36.13	-0.14	0.31	30628
Paternal Education	4.01	4.09	0.08	0.45	30628
Hindu	0.763	0.765	0.002	0.61	30628
Muslim	0.11	0.103	-0.007	0.33	30628
Female decisionmaker for contraception	0.77	0.69	-0.08	0.33	30628
			p-value, joint χ^2 test = 0.61		
			N=30628 ***pseudo-R^2***=0.0011		

Note: This table compares the families with first-born sons and first-born daughters. The first column reports the indicated covariate mean for families with first-born sons, the second column reports the indicated covariate mean for families with first-born daughters, the third column reports the difference between the first and the second columns, the fourth column shows the *p*-values, which are based on a two-sample *t*-test of difference in means assuming equal variances. The last column shows the number of non-missing observations for each covariate. At the bottom, the *p*-value from the joint χ^2-test is shown. Regression sample size and pseudo-R^2 are shown at the bottom.

Table 3: Effect of a first-born female on the number of pregnancies, births, number of living children and contraceptive usage

Mothers age		Number of Pregnancies (1)	Number of Births (2)	Number of Living Children (3)	Contraceptive Usage (4)	Pregnancy Termination (5)
		FERTILITY OUTCOMES				
Age 15-49 (a)	τ_{OLS}	0.218***	0.209***	0.196***	-0.036***	0.003
		(0.025)	(0.019)	(0.017)	(0.002)	(0.005)
	τ_{MLE}	0.071***	0.053***	0.069***		
		(0.009)	(0.006)	(0.004)		
	$\bar{y}\|z_i = 0$	3.72	3.18	2.95	0.83	0.18
Age 15-29 (b)	τ_{OLS}	0.106***	0.095***	0.074***	-0.047***	0.007
		(0.018)	(0.017)	(0.012)	(0.006)	(0.002)
	τ_{MLE}	0.097***	0.045***	0.038***		
		(0.008)	(0.004)	(0.003)		
	$\bar{y}\|z_i = 0$	2.39	2.13	1.98	0.67	0.05
Age 30-39 (c)	τ_{OLS}	0.268***	0.253***	0.245***	-0.023***	0.015*
		(0.029)	(0.019)	(0.017)	(0.004)	(0.003)
	τ_{MLE}	0.074***	0.082***	0.087***		
		(0.006)	(0.008)	(0.011)		
	$\bar{y}\|z_i = 0$	3.4	2.8	2.3		0.21
Age 40-49 (d)	τ_{OLS}	0.288***	0.242***	0.240***	-0.016*	0.013*
		(0.036)	(0.03)	(0.027)	(0.014)	(0.005)
	τ_{MLE}	0.058***	0.047***	0.058***		
		(0.060)	(0.062)	(0.063)		
	$\bar{y}\|z_i = 0$	4.98	4.06	3.94		0.19

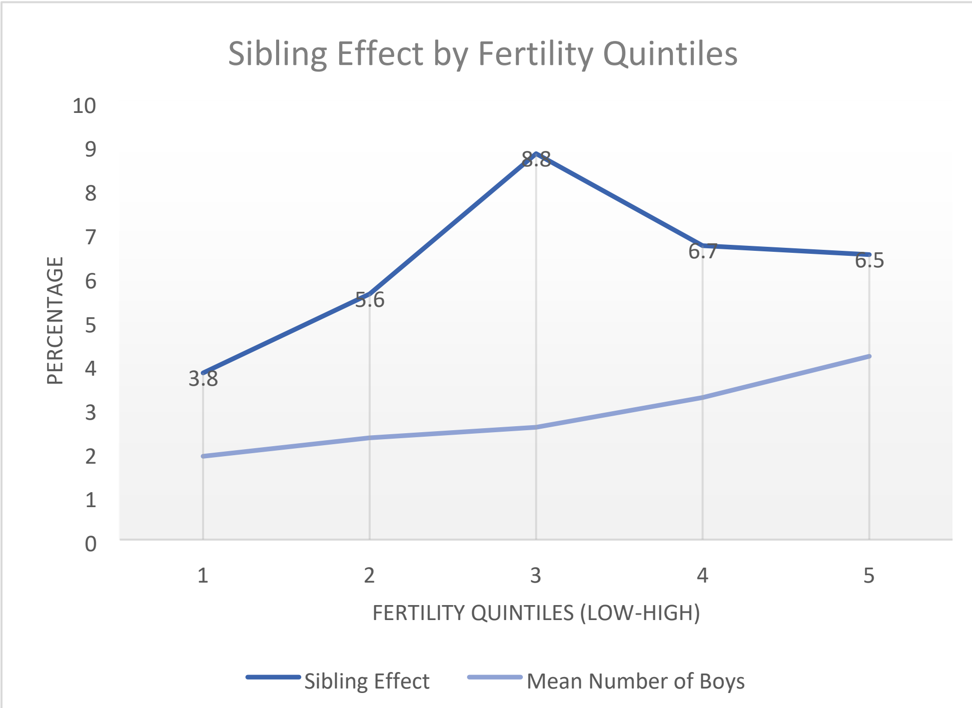
Note: Each column shows the effect of a first-born female on the number of pregnancies, number of births, number of living children, current contraceptive use (includes withdrawal, periodic abstinence, vaginal douche, the pill, injections, female or male condom, intrauterine device, or sterilization), and any pregnancy termination in the past (includes miscarriages, abortions or still births). In each of the panels (A) through (D), for women in the indicated age group, τ_{OLS} shows the OLS estimate and τ_{MLE} shows the maximum likelihood estimate assuming a Poisson process. Mean outcomes for families with first-born males are shown as $\bar{y}|z_i = 0$. All regressions control for the survival of the first-born, year of survey, time specific region fixed effects, mother's age, age at first marriage, years of education, religion, caste, rural residence, husband's age and years of education, husband's years of education. Heteroskedasticity-consistent standard errors are in parentheses. Significance levels are indicated by * < 0.10, ** < 0.05, *** < 0.01.

Table 4: Endogenous Stratification Results on the Number of Living Children

	Repeated Split Sample		***Leave-one-out***				
Fertility Quantile	***Unadjusted (1)***	***Adjusted (2)***	***Unadjusted (3)***	***Adjusted (4)***	$y\|z_i = \mathbf{0}$ ***(5)***	%Δ ***(6)***	N_k ***(7)***
					1.95		
τ_1	**0.08*****	**0.069*****	**0.078*****	**0.073*****		**0.038**	**6673**
	(0.016)	(0.015)	(0.016)	(0.015)			
τ_2	**0.141*****	**0.136*****	**0.157*****	**0.132*****	**2.32**	**0.056**	**6662**
	(0.026)	(0.024)	(0.028)	(0.023)			
τ_3	**0.216*****	**0.205*****	**0.224*****	**0.226*****	**2.56**	**0.088**	**6671**
	(0.037)	(0.023)	(0.042)	(0.025)			
τ_4	**0.225*****	**0.212*****	**0.217*****	**0.219*****	**3.24**	**0.067**	**6673**
	(0.052)	(0.044)	(0.043)	(0.041)			
τ_5	**0.264*****	**0.268*****	**0.286*****	**0.274*****	**4.18**	**0.065**	**6669**
	(0.068)	(0.039)	(0.074)	(0.036)			

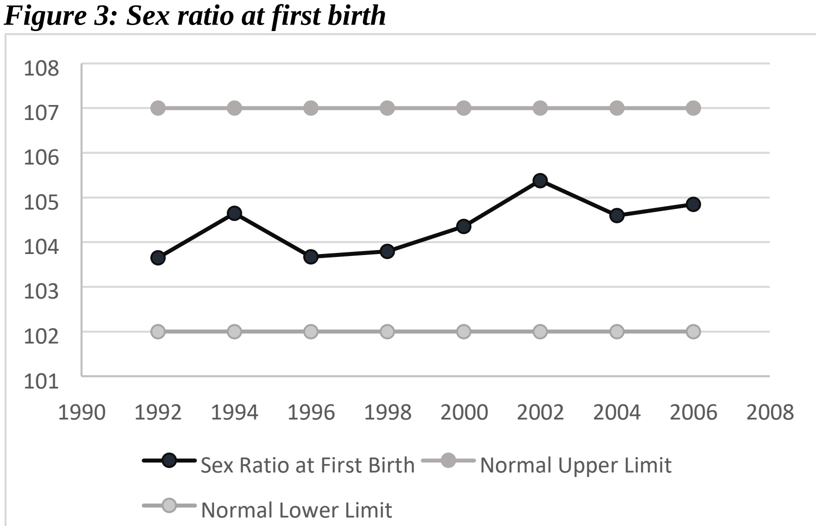
Note: This table shows the effects of a first-born daughter on the number of living children for each of the predicted fertility quantiles. The outcome is the number of living children in the family. Columns (1)-(4) show the treatment effects for each fertility quantile, $\boldsymbol{\tau_k}$ where k = 1,2..5.. Columns (1) and (2) are estimated with the repeated split sample estimator. Columns (3) and (4) are estimated with the leave-one-out estimator. Both estimation methods are provided in Abadie et al. (2014). Column (5) shows the mean number of children for families with a first-born male, indicated with $\bar{y}|\mathbf{z_i} = \mathbf{0}$ for each fertility quantile. Column (6) shows the percentage change (%D) in family size induced by a first-born female and calculated by dividing the treatment effect in column (4) by the mean number of children in column (5). Variables that are used to predict the fertility quantiles are the mother's age at first marriage, mother's and father's years of education, rural residence, and region. The adjusted regressions control for the survival of the first-born, year of survey, time specific region fixed effects, mother's age, age at first marriage, years of education, religion, caste, rural residence, husband's age and years of education, husband's years of education. Heteroskedasticity-consistent standard errors are in parentheses. Bootstrapped standard errors are in parentheses. Significance levels are indicated by * < 0.10, ** < 0.05, *** < 0.01.

Figure 2: Sibling effect by fertility quintiles



Note: Figure 2 shows the percentage change (%D) in family size induced by a first-born female and calculated by dividing the treatment effect in column (4) by the mean number of children in column (5) on the y-axis, and the fertility quintiles from lowest to the highest moving from left-to-right in this figure. The smooth line at the bottom measures the mean number of children if the first-child is a boy.

Figure 3: Sex ratio at first birth



Note: Figure 3 presents the male to female sex ratio at birth for the first child. It lies within the "normal" range as Bhalotra (2010) puts it.

Table 5.A:

Infant Mortality (A)				
	First-Born			
	Boy		*Girl*	
	Second-born		*Second-born*	
	Boy	***Girl***	***Boy***	***Girl***
Mean	0.061	0.068	0.059	0.085
Standard Deviation	[0.23]	[0.29]	[0.23]	[0.22]
Sibling Effect	**0.007**		**0.026*** **	
	[0.006]		[0.005]	
diffs-in-diffs	**0.019*** ** [0.003]			
Adjusted diffs-in-diffs regression results	**0.020*** ** [0.004]			

Table 5.B:

Child Mortality (B)				
	First-Born			
	Boy		*Girl*	
	Second-born		*Second-born*	
	Boy	***Girl***	***Boy***	***Girl***
Mean	0.021	0.019	0.02	0.022
Standard Deviation	[0.23]	[0.29]	[0.23]	[0.24]
Girl-boy difference	**-0.002**		**0.020*** **	
	[0.019]		[0.012]	
diffs-in-diffs	**0.022*** ** [0.017			
Adjusted diffs-in-diffs regression results	**0.017*** ** [0.014]			

Note: These tables compare the health outcomes of the second-born children by first-born sibling's sex. Panel (A) compares the infant mortality rates, Panel (2) compares the child mortality rates. Infant mortality is defined as the death of a child under the age of one. Child mortality is defined as the death of a child under the age of five. Girl-boy difference estimator shows the gender difference in infant mortality by previous sibling's sex. Difference-in-difference estimator shows the difference in girl-boy differences between children who have a previous female sibling and children who have a previous male sibling. The covariate adjusted results are from the regressions that control for the survival of the first-born, year of survey, time specific region fixed effects, mother's age, age at first marriage, years of education, religion, caste, rural residence, husband's age and years of education, husband's years of education. Heteroskedasticity-consistent standard errors are in parentheses. Bootstrapped standard errors are in parentheses. Significance levels are indicated by * < 0.10, ** < 0.05, *** < 0.01.

Appendix Tables on Son Preference and Contraceptive Usage:

Table A1: Ideal number of boys and girls Preferred by households

	Ideal no. of boys (1)	*Ideal no. of girls* (2)	*Ratio - boys:girls* (3)
1992-93	**1.76**	**1.25**	**1.41**
	(0.83)	(0.67)	
1998-99	**1.66**	**1.2**	**1.38**
	(0.78)	(0.62)	
2005-06	**1.52**	**1.16**	**1.31**
	(0.72)	(0.58)	
Religion			
Hindu	**1.49**	**1.07**	**1.39**
	(0.82)	(0.48)	
Muslim	**1.69**	**1.26**	**1.35**
	(0.86)	(0.6)	
Wealth			
Low wealth	**1.39**	**1.19**	**1.16**
	(0.75)	(0.6)	
High wealth	**1.27**	**1.03**	**1.23**
	(0.5)	(0.45)	
Education			
Low Education	**1.57**	**1.18**	**1.33**
	(0.74)	(0.59)	
High education	**1.38**	**1.01**	**1.37**
	(0.46)	(0.44)	

Note: This table is based on a hypothetical question, "if you can start your fertility all over again, what is the number of ideal sons or daughters that you would like to have?" based on NFHS data estimates. Column (1) reports a mean of ideal number of boys households would like to have. Column (2) reports a mean of the ideal number of girls that households would like to have. Column (3) is the ratio of ideal number of boys and girls, derived by dividing column(1) by Column (2).

Table A2: Probit regression estimates for those who desire more children and are not using contraception

Desire more children	***(2005-06)***	***(1998-99)***	***(1992-93)***
Gender			
Son Preference	**0.1105****	**0.0019*****	**0.0118**
	(0.0446)	(0.0002)	(0.0036)
Ratio of sons among living children	**-0.3662*****	**-0.0923*****	**-0.0908*****
	(0.0256)	(0.0155)	(0.0063)
Last child male	**-0.0897*****	**-0.1597*****	**-0.1489*****
	(0.0157)	(0.0127)	(0.0054)
Area			
Urban	**-0.2024*****	**-0.1808*****	**-0.2018*****
	(0.0144)	(0.0132)	(0.0057)
Education			
Maternal education	**-0.0132*****	**-0.0137*****	**-0.0088*****
	(0.0015)	(0.0016)	(0.0005)
Paternal education	**0.0009**	**-0.0735*****	**-0.0944*****
	(0.0007)	(0.0063)	(0.0024)
Religion			
Hindu	**-0.0794*****	**0.0807*****	**0.1109*****
	(0.0307)	(0.0302)	(0.0117)
Muslim	**0.2970*****	**0.4302*****	**0.5337*****
	(0.0344)	(0.0336)	(0.0132)
Christian	**0.2447*****	**0.3134*****	**0.3099*****
	(0.0465)	(0.0432)	(0.0168)
State			
BIMARU	**0.4436*****	**0.6257*****	**0.6196*****
	(0.0144)	(0.0127)	(0.0055)
South	**-0.1576*****	**0.0012**	**-0.0400****
	(0.0173)	(0.0148)	(0.0064)
_Cons	**-0.0863****	**-0.0572***	**0.0707****
	(0.0404)	(0.0316)	(0.0131)

Note: Baseline equation is $F_i = \alpha + \beta_1 SP_i + \beta_2 LC_i + \beta_3 l_i + \gamma_1 u_i + \gamma_3 e_i + \gamma_4 r_i + \gamma_5 s_i + \varepsilon_i$, where, F stands for Fertility Preferences (do not desire any more children) at each cross section i representing the surveys NFHS 1, NFHS 2 and NFHS 3. There are three variables used to measure gender son preference in India, being *SP or* Son Preference taken as the ratio of ideal number of boys desired to Ideal number of children desired; LC being the Ratio of living sons to the total living children; l, used to measure the probability of not using contraception when last child is male; u standing for Urban Residence; e for Education (Maternal, Paternal), r representing Religion (Hindu, Muslim) and S representing the state bunches BiMaRU and South India. BiMaRU is representative of four North Indian states, Bihar, Madhya Pradesh, Rajasthan and Uttar Pradesh. The name BiMaRU has been coined by Ashish Bose in the 1980s, and translates to being ‘ill’ in the Hindi, as this region ranks the least in terms of Gross Domestic Product and Human Development Indicators, while the South Indian states are consistently ranked as the highest in both respects. Significance levels are indicated by * < 0.10, ** < 0.05, *** < 0.01.