Missing Women: A Quantitative Analysis*

Yongkun Yin

CEMFI

Job Market Paper

October 2022

Please click here for latest version

Male-biased sex ratios (missing women) and gender education gaps in some developing countries have raised much concern. Why do parents favor sons over daughters? What policies are effective in normalizing the sex ratio? To answer these questions, an overlapping-generation model of fertility, sex selection, the quantity-quality trade-off, and marriage is built and estimated for India. The quantitative analysis reveals that the main drivers of missing women are economic factors, i.e., old-age support by sons, dowry payment for daughters, and labor market discrimination against women. If the gender differences in the economic factors are removed, the male-female ratio at birth (or sex ratio at birth, SRB) would reduce from 1.139 to 1.052. The fertility rate would drop from 3.07 to 2.40, and the share of women with secondary education would increase from 49.2% to 72.1%. Once the economic factors become symmetric for sons and daughters, eliminating intrinsic son preference has a small additional effect. Prohibiting prenatal sex selection could normalize the sex ratio, but it lowers educational attainment for women. A subsidy for female births or a subsidy for female education can reduce the SRB. But the former increases fertility and reduces children's education and women's labor supply, while the latter has opposite effects. Finally, a pay-as-you-go pension system can lower the SRB to 1.092, but it also reduces children's educational attainment, as parents value them now less.

Keywords: Son preference; Sex selection; Fertility; Education; India.

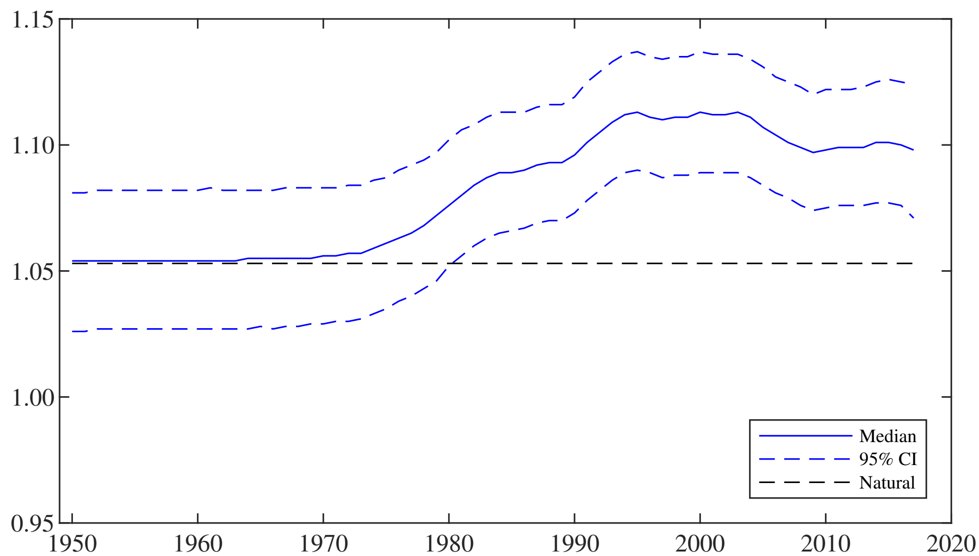
JEL codes: D1, J13, O41.

---

*Yongkun Yin, CEMFI, Madrid, Spain. Email: yongkunyin@gmail.com. I am grateful to Nezih Guner for his support and guidance at different stages of the project. I also thank Sonia Bhalotra, Samuel Bentolila, Sebastián Fanelli, Raquel Fernández, Pedro Mira, Josep Pijoan-Mas, Alessandro Ruggieri, and seminar participants at CEMFI for insightful comments and discussions. I acknowledge the funding from Fundación Ramón Areces.

1 Introduction

In several developing countries, the ratio of men to women is way too high, a situation that Sen (1990) famously called "missing women." Figure 1 shows the male-to-female ratio at birth (or sex ratio at birth, SRB) in India since the 1950s. The natural value for the SRB is around 1.05 and is quite constant across time and space due to biological reasons (Orzack et al., 2015). The SRB in India began to increase in the 1980s with the spread of ultrasound machines, which can be used for prenatal sex detection. In the 1990s, the ratio exceeded 110 male births per 100 female births, well above the natural balance.



Notes. Data are from Chao et al. (2019), who estimate the SRB and the natural ratio using a Bayesian method for each country.

Figure 1. Ratio of Male to Female Births in India.

A biased sex ratio at birth or son preference is not unique to India. Parents in different parts of the developing world want to have more sons than daughters and care more about sons' education than daughters' (Williamson, 1976; Fuse, 2010).[1] In China, India, and other (mainly South and East Asian) countries, parents' desire to have a son results in sex-selective abortions and excessive female child mortality (e.g., Hesketh and Zhu, 2006; Dyson, 2012). Other social problems have emerged, such as the low education levels

[1]Most studies focus on developing countries. However, evidence has also been documented for currently developed countries, such as Germany (Sandström and Vikström, 2015), Greece (Beltrán Tapia and Raftakis, 2021), Spain (Marco-Gracia and Beltrán Tapia, 2021) and the US (Dahl and Moretti, 2008; Blau et al., 2020).

of women (Barcellos et al., 2014) and the shortage of women in the marriage market (Guilmoto, 2012; Jiang et al., 2014).

Why are there missing women? Most studies attribute parental preferences for sons to the gendered benefits and costs of children, such as old-age support provided primarily by sons (Das Gupta et al., 2003; Chung and Das Gupta, 2007) and dowries payment for daughters (Diamond-Smith et al., 2008; Bhalotra et al., 2020). The value of daughters may also be reduced by their disadvantaged labor market conditions (Qian, 2008; Carranza, 2014). In addition, a preference for sons can be a feature of parental utility function, known as intrinsic son preference.[2] Despite these insights, surprisingly little is known about the relative importance of these factors, how these factors affect the aggregate economy, and what policies can be most effective in correcting the biased sex ratio and its associated problems. The current paper tries to fill this gap.

Son preference may have important effects on fertility and children's education. Since parents tend to have another birth if they have no son (Clark, 2000; Altindag, 2016), it can increase the family size and lower education investment per child (Arnold, 1985; Bairagi and Langsten, 1986; Leung, 1994; Dharmalingam et al., 2014).[3] Parents may also shift resources directly from daughters to sons (Barcellos et al., 2014; Choi and Hwang, 2015).

Fertility and education may also be affected by sex selection technologies (i.e., prenatal sex detection and abortion technologies), which have been spreading in developing countries for decades. First, sex selection technologies can alter the quantity-quality trade-off of parents. On the one hand, to achieve the ideal number of sons, they can directly select male births instead of practicing a son-biased fertility stopping rule. This may reduce the family size and increase education investment per child. On the other hand, since a birth of the selected sex becomes more valuable, parents may want more births. Second, sex se-

[2]There are a few reasons for intrinsic sex preference. Evolutionary biology suggests that human beings prefer the sex with higher fitness (Trivers and Willard, 1973). In patrilineal societies, sons take the responsibilities to continue the family line and perform ancestor worship to ensure the welfare of deceased parents (Das Gupta et al., 2003; Pande and Astone, 2007). In addition, parents may develop intrinsic sex preference after being exposed to such norms for a long time (Almond et al., 2013; Carol and Hank, 2020).

[3]To estimate the effect of sex preference on fertility, the literature often uses the method proposed by Arnold (1985), which first computes the parity progression ratios conditional on the sex composition of previous children and then computes a hypothetical fertility rate supposing that all couples at each parity are as likely to have another birth as those at the same parity but least likely to have another birth.

lection technologies can change the trade-off between prenatal sex selection and postnatal sex discrimination. Girls born despite the accessibility of the sex selection technologies are more "wanted", so they are expected to receive more investment (Goodkind, 1996; Davies and Zhang, 1997; Anukriti et al., 2022 Rastogi and Sharma, 2022).

In this study, I focus on India and ask the following questions: What are the relative contributions of different factors to the imbalanced sex ratio and gender education gap? In particular, what is the relative importance of the economic factors, i.e., old-age support requirements, dowries, and labor market discrimination against women, versus intrinsic son preference? How can the factors driving missing women affect fertility and female labor supply in equilibrium? What are the impacts of sex selection technologies? What policies can normalize the sex ratio, and how can the policies affect fertility and female labor supply?

To answer these questions, a quantitative overlapping-generation model embedding the standard quantity-quality trade-off and marriage is built, which incorporates various causes of missing women and a sex selection technology. Parents pay dowries to married daughters and receive transfers from adult children, mainly from their sons, when they get old. There is also labor market discrimination, which acts as a wedge between women's productivity and wage. These economic factors could affect parental preference for the quantity and quality of sons and daughters. In addition, parents have an intrinsic sex preference. They derive utilities directly from the number of children and their education investment, but they value daughters less than sons. Fertility decisions are made sequentially conditional on the sex composition of the previous children. When having children, parents can use a sex selection technology to ensure a male or female birth at a monetary and a utility cost.

The model is calibrated to the Indian data. It does a good job replicating the household behaviors, including savings rate, women's working hours, women's education-fertility gradient, education expenditure per son and per daughter, the distribution of men and women's education, and the probabilities of having another birth and the likelihood that the next birth is a boy conditional on the sex composition of the previous children in the data. As a validation exercise, the model can also replicate the quasi-experimental evidence associated with policy interventions in India (Biswas et al., 2021).

A series of experiments are conducted to quantify the relative contribution of different factors to missing women and their effects on fertility, education, and female labor supply. I start by removing the gender difference in one economic factor at a time and then remove all of them simultaneously. Next, I eliminate intrinsic sex preference.

The first finding is that economic factors play a relatively more important role than intrinsic sex preference. In the benchmark model, the SRB is 1.139. 64.6% of men complete secondary education, but only 49.2% of women achieve this. If there is no gender difference in the economic factors, the SRB will drop to 1.052, which is very close to the natural ratio of 1.05. The share of men with secondary education would rise to 79.2%, while the share of women would rise more significantly to 72.1%. If the intrinsic sex preference is also eliminated, the SRB would drop further, but only by 0.002 to 1.05. The share of men with secondary education would fall slightly to 77.9%, but the share of women with secondary education would rise slightly to 77.4%.

The second finding is that the economic factors have much more significant effects than intrinsic son preference on other outcomes of interest, including fertility and female labor supply. In the benchmark model, the fertility rate is 3.07, young women supply 25.0% of their time to the market, and middle-aged women supply 27.0%. If there is no gender difference in the economic factors, the fertility rate will drop by 0.67 to 2.40, and women would supply 44.8% and 45.1% of their time to the market when they are young and middle-aged, respectively. If intrinsic son preference is further eliminated, the marginal effects are minimal. The fertility rate would further drop by 0.06. Young and middle-aged women would supply 45.9% and 46.1% of their time to the market, respectively.

The third finding is that labor market discrimination alone is not a driving force for sex selection. If only labor market discrimination is removed from the benchmark economy, the SRB will increase to 1.232 instead of decreasing. The reasons are twofold. First, the value of daughters does not increase as long as sons are the primary providers of support for parents, and the value of sons increases even more now since their wives contribute more to household income. Second, women supply more labor to the market and have fewer children, which raises the likelihood of no son if they do not abort female fetuses. Both factors interact and generate incentives for parents to select a male birth.

What happens if sex selection is prohibited? Although such a policy can normalize the SRB by force, it reduces education expenditure per daughter by 1.0% and reduces the share of women completing secondary education by 1.0 pp. There is also a decline in the share of men with secondary education by 1.0 pp due to a small increase in the fertility rate.

Finally, policies that can potentially normalize the sex ratio are explored. I start by comparing a subsidy for female births with a subsidy for female education, two measures that Indian governments widely use. The analysis reveals that although both policies can reduce the sex ratio, they have different effects on female education, fertility, and female labor supply because they shift the parental quantity-quality trade-off into different directions.

By lowering the fixed cost of each birth, the subsidy for female births increases fertility, which reduces education expenditure per child and decreases women's labor supply. Suppose the subsidy amounts to 40% of the annual income of a male without secondary education. It would reduce the SRB to 1.063 and raise the fertility rate to 3.55. The share of men and women with secondary education would drop to 54.0% and 38.5%, respectively. Finally, the labor supply by young women would drop to 20.7% of their time, while it rises slightly to 28.4% for middle-aged women. In contrast, by motivating parents to invest more in girls' education, female education subsidy reduces fertility, further increasing women's labor supply. The results suggest that a subsidy rate of 20% would reduce the SRB to 1.091 and the fertility rate to 2.82. The share of women with secondary education would go up significantly to 56.8%, while the share of men with secondary education would increase slightly to 65.0%. Finally, women would supply 30.5% of their time to the market when they are young and 31.5% when they are middle-aged.

Next, the impacts of a pension system are assessed. As another source of income after retirement, pension weakens the role of old-age support provided by children and the gender difference in old-age support. Therefore, fertility, education expenditure, and the sex ratio are reduced. Suppose that the pension system is pay-as-you-go with a replacement rate of 50%. This is the targeted rate of the Employees Provident Fund Scheme, the largest pension scheme in India. The SRB would drop to 1.092, fertility would decrease by 0.04 to 3.03, the share of men with secondary education would fall by 8.0 pp to 56.6%, and the share of women with secondary education would fall by 4.8 pp to

44.4%. If the pension system reduces the old-age support requirement for children, the sex ratio will drop more significantly. With old-age support responsibilities reduced by half, the SRB drops to 1.084. While the share of men with secondary education further drops to 55.3%, the share of women with secondary education rises slightly to 44.9% compared with the case without any change in old-age support. The reason for this improvement is the reduction in old-age support motive for educating sons, which leads parents to shift more resources to daughters due to altruism.

The paper builds on the macro-development literature on fertility and investment in children. Many studies investigate the joint evolution of fertility and economic development driven by some third factors, such as productivity growth (Galor and Weil, 2000; Greenwood and Seshadri, 2002; Hansen and Prescott, 2002; Delventhal et al., 2021; Adams, 2021) and mortality decline (Kalemli-Ozcan, 2002, 2003; Bar and Leukhina, 2010). In this study, I investigate how the factors driving missing women can affect fertility and human capital. It is also related to recent studies examining how fertility-related technologies affect fertility and human capital. In particular, Cavalcanti et al. (2021) investigate how family planning interventions subsidizing contraceptive use can improve living standards by reducing unwanted fertility, while Seshadri and Zhou (2022) explore how family planning interventions could improve intergenerational mobility by reducing unwanted fertility among the poor. In this study, I study how sex selection technologies could affect the number of sons and daughters and their human capital.

This paper is also related to the economic and demographic literature on the causes of biased sex ratio. Some studies attribute it to the patrilocal family arrangement, whereby sons provide more financial support to old parents than daughters (e.g., Das Gupta et al., 2003; Chung and Das Gupta, 2007; Ebenstein and Leung, 2010). Others argue that dowry payments shape parental preference toward fewer daughters (Diamond-Smith et al., 2008; Bhalotra et al., 2020). Moreover, disadvantaged labor market opportunities for women restrict the value of daughters accrued to parents (Boserup, 1970; Bardhan, 1974; Rosenzweig and Schultz, 1982; Qian, 2008; Carranza, 2014). Finally, many studies show that migrants from countries with son preference to Western countries still prefer sons to daughters even if the socioeconomic environments have changed, pointing to the role of intrinsic son preference (e.g., Dubuc and Coleman, 2007; Abrevaya, 2009; Almond et al., 2013). Unlike these studies, which may focus on one factor, my study uses a quantitative model to quantify the relative importance of these factors. Therefore, it sheds light on possible policies that can normalize biased sex ratios.

Finally, it is related to the large empirical literature evaluating the impacts of various policies on sex ratio, fertility, and investment in children.[4] In particular, Nandi and Deolalikar, 2013 examine the effects of prohibiting sex selection on the sex ratio at birth, and Rastogi and Sharma (2022) further investigate how this policy affects the education of girls. There are also studies evaluating conditional cash transfer programs targeted at families with daughters (Sinha and Yoong, 2009; Powell-Jackson et al., 2015; Anukriti, 2018; Biswas et al., 2021). In general, such studies focus on the partial equilibrium effects of a small-scale policy. A quantitative model allows me to assess the impact at the aggregate scale and examine possible policy interactions.

The remainder of the paper is organized as follows. Section 2 documents some empirical facts that motivate this study. Section 3 describes the model, which will be used for the quantitative analysis. Section 4 explains how the parameters are calibrated. Section 5 assesses the causes and consequences of missing women. Section 6 examines possible policies to reduce imbalanced sex ratios and the gender education gap. Section 7 concludes.

2 Facts

In this section, I describe the empirical facts that motivate this work. First, I show how fertility and sex selection depend on the sex composition of current children. Next, I document the negative relationship between the quantity and the quality (education) of children and study how fertility differs for parents with different education.

Data are from the Indian Demographic and Health Survey (Indian DHS) and the Indian Human Development Survey (Indian HDS). The DHS is a repeated cross-sectional survey collecting high-quality representative data on population, health, and nutrition in developing countries. In India, it was conducted in 1992-1993, 1997-1998, 2005-2006, and 2015-2016. In this survey, women aged 15-49 reported their complete fertility history,

[4] Kumar and Sinha (2020) provide a review of the policies aimed at normalizing the sex ratio and studies evaluating such policies in India and some other developing countries.

including the sex, date, and survival status of each birth. Based on this information, I can examine household decisions on fertility and sex selection. In the analysis, I focus on women aged 40-49 years who had almost completed their fertility and restrict the sample to those with 1-6 children. The 2015-2016 data are used for the main analysis, and the data from the previous waves are used for robustness checks.

The Indian HDS is a nationally representative, multi-topic longitudinal survey. The first wave was conducted in 2004-2005, and the second wave was in 2011-2012. In the second wave, married women reported the number of years of education, the number of male and female siblings, and parental education for both herself and her husband. Based on this information, I can examine the relationship between the education of the wife/husband and the number of children in her/his natal family.

Fact 1. Son-biased Fertility Stopping Rule.

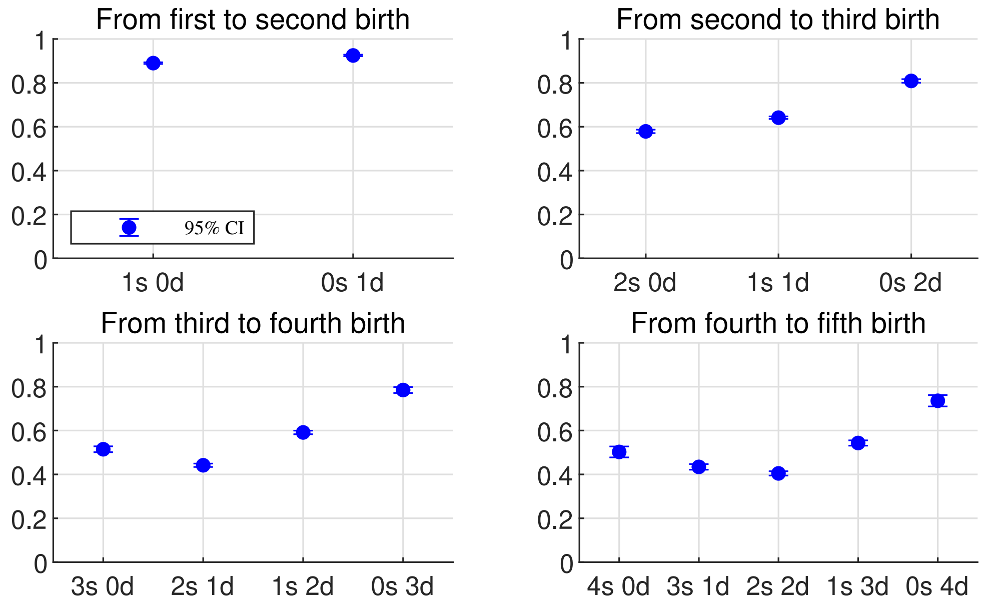
The Indian DHS data in 2015-2016 suggest that Indian parents want about 47.5% of their children to be boys but only 40.5% to be girls, suggesting a strong favor for sons.[5] What will parents do if they want sons but do not have (enough) sons? A typical practice is to continue childbearing until reaching their ideal number of sons (e.g., Clark, 2000; Altindag, 2016). This subsection examines how parental decision on whether to have another birth depends on the sex of existing children. Based on data from the Indian DHS in 2015-2016, I compute the parity progression ratio (PPR), or the proportion of women who progressed from one parity to the next, conditional on the sex composition of existing children.[6] The results are plotted in Figure 2.

Two patterns emerge. First, parents are more likely to have another birth if their children have an imbalanced sex composition, i.e., if they either have too few sons or too few daughters. The pattern is most salient among women with three or four children.[7]

[5] Parents were asked about the number of children they would like to have if they could go back to the time without any children and the number of boys and girls they would like to have. Based on the answers, I compute the ideal proportion of sons, the ideal proportion of daughters, and the proportion of children whose sex does not matter. On average, the ideal proportion of sons is 47.5% for both men and women; the ideal proportion of daughters is 41.4% for women and 39.4% for men.

[6] It should be note that respondents might have recall bias since the data are retrospective. Anticipating this problem, the DHS programs used numerous probes to encourage recall accuracy. However, recall bias might still exist for births that occurred many years ago (Bhalotra and Cochrane, 2010; Rosenblum, 2013).

[7] This pattern cannot be observed for women with two children, as the progression ratio is lower for women with two daughters than for women with a son and a daughter. As will be clear soon, parents not only want a balanced sex composition of children but also want more sons than daughters. For women



Notes. Data are from the Indian Demographic and Health Survey in 2015-2016.

Figure 2. Parity Progression Ratio Conditional on the Sex Composition of existing Children.

Consider women with four children for example, if they have two sons and two daughters, only 40.5% of them will have another birth. However, if they have only daughters or only sons, the ratio increases sharply to 50.3% and 73.6%, respectively. The U shape of the PPR suggests that parents do not treat sons and daughters as perfect substitutes and that they want a balanced sex composition of children.

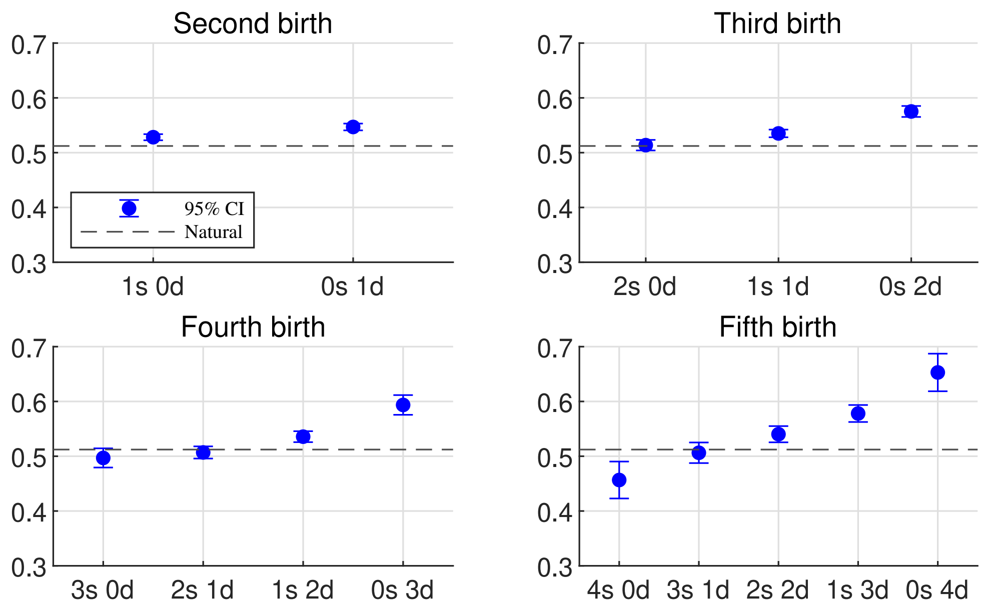
Second, parents are more likely to have another birth if sons make up a small proportion of existing children than if daughters make up a small proportion. That is, parents practice a son-biased fertility stopping rule. For example, consider women with two children, 57.9% of them will have another birth if they have two sons and no daughter, but 80.9% of them will have another birth if they have no son and two daughters. The asymmetry in the PPR suggests that parents have a preference for sons and want more sons than daughters.

Fact 2. Son-biased Sex Selection.

After the middle 1980s, prenatal sex detection was increasingly accessible in India with the spread of ultrasound machines. Parents began to use abortion to select the sex of their children (Bhalotra and Cochrane, 2010). This subsection examines how the

with two children, the preference for sons dominates the preference for a balanced sex composition.

parental decision on whether to select the sex of the next birth is affected by the sex of existing children. Based again on data from the Indian DHS in 2015-2016, I compute the probability that the next birth is a male conditional on the sex composition of existing children, which is plotted in Figure 3.



Notes. Data are from the Indian Demographic and Health Survey in 2015-2016. The black dashed line indicates the natural level.

Figure 3. Probability that the Next Birth is a Boy Conditional on the Sex Composition of Previous Children.

A striking finding is that the next birth is more likely to be a boy for parents who have a smaller proportion of sons. This pattern can be observed at all the birth parities. More specifically, if sons make up a large proportion of existing children, the probability that the next birth is a boy will be very close to the natural probability of around 51.2%. However, if parents have too few sons, the probability will be significantly higher than the natural probability. Consider women who have a third birth for example, if the first two children are both boys, 51.4% of the third births will be boys. However, if the first two children consist of a boy and a girl, 53.5% of the third births will be boys, 2.5 pp higher than the natural probability. Furthermore, if the first two children are both girls, 57.5% of the third births will be boys, 6.3 pp higher than the natural probability. The deviation from the natural probability provides strong evidence that parents have disproportionally aborted female fetuses to ensure a male birth.

Finally, it is noteworthy that parents may also select a female birth when they have a very small proportion of daughters due to the preference for a balanced sex composition of children. For instance, for women who have a fifth birth after having four sons, 45.7% of the fifth births are a boy, 5.5 pp lower than the natural probability. However, son-biased sex selection is the dominant practice.

Fact 3. Quantity-Quality Trade-off.

To examine the relationship between the quantity and quality of children, I exploit the data from the Indian HDS in 2011-2012. Regressions are run for married males and females, where the dependent variable is the years of schooling of the male/female, and the independent variables are the number of male and female children in his/her natal family. In addition, I control for the father's years of education, the mother's years of education, and the birth year fixed effect of the male/female of concern. I focus on the sample of males and females aged 23-32 in the main analysis and use other age groups for robustness checks.[8] The empirical analysis is conducted for males and females separately, and the results are reported in the first two columns of Table 1, respectively.

Column (1) reveals that as the number of boys and girls in a family increases, boys receive fewer years of school education. More specifically, one more boy in the family reduces the education of boys by 0.54 years and one more girl reduces it by 0.37 years. Column (2) reveals a similar pattern for the education of girls. One more boy reduces the education of girls by 0.39 years and one more girl reduces it by 0.11 years.

One may also notice that compared with boys, girls' education is less affected by the number of boys and girls in the family. In addition, children's education is less affected by the number of girls than by the number of boys in the family. One possible reason is that most education resources are allocated among boys rather than girls. As will be shown next, girls receive less education than boys.

Fact 4. Gender Education Gap.

To examine the gender education gap, I exploit the Indian HDS data in 2011-2012 again. The samples of married males and females used to document Fact 3 are pooled

[8] Individuals attending college are expected to graduate at age 22. So I use the sample of individuals aged above 22. I focus on those aged 32 and below because the socioeconomic environments, which affect household decisions, can be different for individuals born in different years.

Table 1: Family Size, Gender and Years of Schooling

	(1) Males	(2) Females	(3) Both
Number of male children	-0.544***	-0.390***	-0.470***
	(0.043)	(0.050)	(0.034)
Number of female children	-0.366***	-0.109**	-0.262***
	(0.035)	(0.044)	(0.028)
Male			1.939***
			(0.085)
Father's years of schooling	0.349***	0.408***	0.371***
	(0.013)	(0.014)	(0.010)
Mother's years of schooling	0.436***	0.211***	0.355***
	(0.016)	(0.017)	(0.012)
Birth year fixed effects	Yes	Yes	Yes
N	12159	8415	20574
R^2	0.381	0.289	0.354

Notes. Sample consists of men and women aged 23-32 years. Standard errors in parentheses. *** $p < 0.01$.

together. The regression is repeated by adding a dummy variable indicating whether the child is a male or not. The results are reported in Column (3) of Table 1, which reveal that female children receive about two fewer years of education than male children, controlling for the number of boys and girls in the family.

Fact 5. Women's Education and Fertility.

Another fact closely related to the quantity-quality trade-off is that parents with more education generally have fewer children but invest more in children's education. To examine the relationship between education and fertility, I use the data from the Indian DHS in 2015-2016. A regression is run, where the dependent variable is the number of children ever born, and the independent variables of interest are the years of education of the female and her spouse. Meanwhile, both spouses' birth year fixed effects are controlled for.

The results in Table 2 suggest that parental education reduces fertility and women's education has a particularly strong effect. More specifically, if the husband receives ten more years of education, they will have 0.25 fewer children, while if the wife receives ten

Table 2: Education and Fertility

	(1)
	Children ever born
Wife's years of schooling	-0.115***
	(0.003)
Husband's years of schooling	-0.025***
	(0.003)
Wife's birth year fixed effects	Yes
Husband's birth year fixed effects	Yes
N	22469
R^2	0.152

Notes. Sample consists women aged 40-49 years. Standard errors in parentheses. *** $p < 0.01$.

more years of education, they will have 1.15 fewer children.

Robustness Checks

Extensive robustness checks are conducted for the facts documented above. The analysis is repeated for different groups of people in the same data, using different waves of the same survey, or for different countries. The results are summarized in brief below and described in detail in Appendix A.

Fact 1 is based on the data from the Indian DHS in 2015-2016. (i) It can be observed for women with and without secondary education. (ii) It can also be observed in the previous waves of the survey in 1992-1993, 1998-1999, and 2005-2006. (iii) It can also be observed in Nepal and Vietnam, where survey evidence indicates that women want more sons than daughters (Fuse, 2010).

Fact 2 is based on the data from the Indian DHS in 2015-2016. (i) It can be observed for both women with and without secondary education. (ii) It can also be observed in the previous waves of the survey in 1992-1993, 1998-1999, and 2005-2006, although parents in these waves were less likely to select their children's sex. (iii) It can also be observed in Nepal and Vietnam.

Fact 3 and Fact 4 are based on the sample of men and women aged 23-32 in the Indian HDS in 2011-2012. It can also be observed for individuals aged 33-42 and 43-52.

Fact 5 is based on the Indian DHS in 2015-2016. (i) It can also be observed in the previous waves of the survey in 1992-1993, 1998-1999, and 2005-2006. (ii) It can also be observed in Nepal and Vietnam.

3 The Model

This section builds an overlapping-generation model of fertility, investment in children, and marriage formation. Individuals in the model economy live for four periods: period 0 as children, period 1 as young adults, period 2 as middle-aged adults, and period 3 as old adults. There are two genders (sexes), male and female, $g \in \{f, m\}$. Children do not make any economic decisions. They receive education and acquire human capital. Individuals can have two levels of education/human capital: low and high, $h_g \in \{h_{g,L}, h_{g,H}\}$, determined by parental education investment. Once children become young adults, they enter the marriage market. Newly-wed couples receive a dowry from the parents of the bride. Young married couples have children, invest in children's education, work, and consume. Husbands only work in the labor market, while wives split their time between market work, household production, and childcare. Parents can select their children's sex, at a monetary and utility cost.

Middle-aged adults work, consume, and save for their old age. They also pay dowries to their young married daughters, if they have any. In addition, middle-aged households make transfers to their old parents. These transfers depend on the number of siblings of the husbands and wives and their gender composition. Old adults do not work and simply consume their savings and transfers from their children. Next, I describe the model in detail.

3.1 Households

Young Married Households

Consider a young couple with the human capital levels $\{h_m, h_f\}$. The husband is assumed to have one unit of time, which is supplied inelastically to the market. The wife also has one unit of time, which can be allocated between childbearing, home production, and market work (l_1). Upon marriage, the young couple receives a dowry (d) from the wife's parents. The household receives a lump-sum transfer T, which is the redistribution of the distortionary costs of labor market discrimination. The household income is,

$$I_1 = h_m + \lambda h_f l_1 + d + T, \tag{1}$$

where $\lambda \in [0, 1]$ captures the labor market discrimination against women.[9]

Starting with 0 children, a young couple decides whether to give a birth, $b \in \{0, 1\}$. For each birth, parents decide whether to use the sex selection technology. Let $s_m \in \{0, 1\}$ and $s_f \in \{0, 1\}$ denote whether the couple uses the sex selection technology to select a male or a female child, respectively. Without selection, each child can be a male with a probability p_m and a female with a probability $1 - p_m$. With sex selection, parents can ensure a male or a female child. But they need to pay a monetary cost π and incur a utility cost ξ. The monetary cost captures what the household has to pay for prenatal sex detection and sex-selective abortions. The utility cost captures the psychological cost of practicing abortion, concern about its side effects, etc., which can be different across individuals.[10] Therefore, I assume that ξ is drawn randomly from a log-normal distribution after a marriage, i.e., $\log(\xi) \sim N(\mu_\xi, \sigma_\xi)$.

After parents have the first child and realize its sex, parents decide whether to have another one. In this way, fertility and sex selection decisions are made sequentially conditional on the sex composition of existing children.

Suppose that parents end up with q_m male children and q_f female children, among which q_s were born with sex selection. Then parents decide how much money to spend on children's education, which can be different for sons and daughters. Denote education expenditure per son by e_m and education expenditure per daughter by e_f. Household consumption of market goods is then given by

$$c_1 = I_1 - q_m e_m - q_f e_f - q_s \pi. \tag{2}$$

Each birth also requires a time commitment of the mother denoted by τ. The time not spent on market work and raising children is used to produce non-market (domestic) goods

[9]Following Cavalcanti and Tavares (2016), I assume that discrimination costs are redistributed in lump-sum back to working households. Therefore, discrimination affects household decisions only because of the wedge between women's productivity and wage but not because of any waste of resources.

[10]For example, Muslims are more reluctant to practice abortions than Hindus, although they have similar levels of reported son preference, measured by their ideal sex composition of children (Bhalotra et al. (2021)).

using a linear production technology. Therefore, household consumption of non-market goods is,

$$n_1 = 1 - l_1 - (q_m + q_f)\tau. \tag{3}$$

Parents are altruistic toward their children. They derive utility from the number of children $\{q_m, q_f\}$ and the education investment of children $\{e_m, e_f\}$. However, they differentiate between sons and daughters. Utility derived from children is,

$$U_a(q_m, q_f, e_m, e_f) = \alpha \log\{[q_m(e_m + e_0)^\eta]^{\frac{\epsilon-1}{\epsilon}} + \omega[q_f(e_f + e_0)^\eta]^{\frac{\epsilon-1}{\epsilon}}\}^{\frac{\epsilon}{\epsilon-1}}. \tag{4}$$

The parameter α represents the level of altruism, which is drawn randomly from a log-normal distribution after the marriage, $\log(\alpha) \sim N(\mu_\alpha, \sigma_\alpha)$. Parental education investment $\{e_m, e_f\}$ is augmented by a constant term e_0, which can be interpreted as public education expenditure.[11] The parameter η governs the relative importance of the quality of children (education investment) compared with the quantity of children. ω indicates the relative importance of daughters compared with sons. If $\omega < 1$, parents have an intrinsic preference for sons, and if $\omega > 1$, an intrinsic preference for daughters. The quantity and the quality of sons and daughters are aggregated using a constant-elasticity-of-substitution (CES) function, and ϵ governs the elasticity of substitution between sons and daughters. If $\epsilon > 1$, sons and daughters are substitutes, and if $\epsilon < 1$, they are complements. If we set $\omega = 1$, $\epsilon = +\infty$, and $e_m = e_f$, then we have a standard functional form common in the literature (e.g., de la Croix and Doepke, 2003; Cavalcanti et al., 2021; Delventhal et al., 2021).[12]

Whether children can achieve a high education is uncertain. The probability that children achieve a high education depends on their education expenditure,

$$Prob(h'_g = h_{g,h}|e_g) = 1 - \exp(-ze_g),\ \ g \in \{m, f\}, \tag{5}$$

where $z > 0$ governs the efficiency of human capital production. The probability of high education is increasing in education expenditure. It goes to 0 as education expenditure goes to 0, and goes to 1 as education expenditure goes to $+\infty$.[13]

[11] An alternative interpretation is that children are endowed with some innate abilities which are valued by parents.

[12] Another standard form is dynastic utility. That is, parents directly care about children's utility (e.g., Barro and Becker, 1989; Carlos Córdoba and Ripoll, 2019). The two functional forms are closely related since education investment for children is positively related to their utility in the future.

[13] The probability of achieving a high education depends on private education expenditure but not

Middle-Aged Married Households

A middle-aged male has one unit of time, which is supplied inelastically to the market. A middle-aged female has one unit of time, which is allocated between market work (l_2) and home production ($n_2 = 1 - l_2$). In addition to labor market earnings, middle-aged households receive a lump-sum transfer T, which is the distributed discrimination cost. The couple chooses savings (a) for their old age.

If parents have any daughters, the daughters grow up and go into the marriage market in this period. Parents need to pay dowries to the newlyweds. Dowries per daughter is an exogenous fraction (δ) of the income of the male.[14,15] As a result, household income net of dowry and savings is,

$$I_2 = h_m(1 - \delta q_f) + \lambda h_f l_2 + T - a. \tag{6}$$

Middle-aged households also need to make monetary transfers to their old-aged parents. The transfer to the husband's parents is an exogenous fraction, t_m, of their income net of dowry and savings (I_2), while for the wife's parents, it is t_f. As will be clear below, t_m and t_f are the functions of the numbers of male and female siblings of the husband and the wife, respectively. Therefore, consumption of market goods by the middle-aged couple is,

$$c_2 = (1 - t_m - t_f)I_2. \tag{7}$$

Old Married Households

Old parents do not work. They consume their savings $a(1 + r)$ and transfers from their children, where r is the interest rate. A son transfers a fraction t_m of his household

public education expenditure. The interpretation is that public education ensures children achieve at least a low education. However, to achieve a high education, parents need to make an additional private investment.

[14]The historical origin of the dowry custom and the determinants of dowry values are not clear (Anderson, 2007). Although dowries are often believed to be a price to clear the marriage market (Becker, 1981), the empirical evidence is mixed. First, while Rao (1993) documents a positive effect of the female-male ratio on dowry values, more recent studies find little evidence (Edlund, 2000; Chiplunkar and Weaver, 2021). Second, education of the bride, a measure of her quality, may increase (Behrman et al., 1999), decrease (Maertens and Chari, 2020; Calvi and Keskar, 2021; Goel and Barua, 2021) or have no effect on (Edlund, 2000) dowries. Finally, education of the groom may increase dowries (Behrman et al., 1999; Maertens and Chari, 2020; Chiplunkar and Weaver, 2021) or have no effect (Edlund, 2000; Calvi and Keskar, 2021). Since dowry is generally positively associated with parental income or wealth (Anderson, 2007; Maertens and Chari, 2020), I assume that dowry is proportional to the father's income.

[15]Assuming dowry payment to be proportional to parental income rather than only the father's income will make the model less tractable without adding any new insights.

income, while a daughter transfers t_f. Both t_m and t_f are functions of the number of sons and daughters, which are given by

$$t_m(q_m, q_f) = \begin{cases} t_{m1} q_m^{-\kappa_m}, & \text{if } q_m > 0, \\ 0, & \text{if } q_m = 0, \end{cases} \tag{8}$$

and,

$$t_f(q_m, q_f) = \begin{cases} 0, & \text{if } q_m > 0, \\ t_{f1} q_f^{-\kappa_f}, & \text{if } q_m = 0. \end{cases} \tag{9}$$

The parameter t_{m1} is the level of transfer when parents have only one son, and $\kappa_m \in (0, 1)$ governs how the transfer per son varies in the number of sons. The transfer from each son is decreasing in the number of sons but the total transfer is increasing in it. Likewise, t_{f1} is the level of transfer when parents have only one daughter, and $\kappa_f \in (0, 1)$ governs how the transfer per daughter varies in the number of daughters.[16]

An important feature of this specification is that daughters do not provide old-age support to their parents unless they have no male siblings. This specification is chosen to capture the observed pattern of old-age support in India. In India, old-aged parents receive support from children by residing with them, and parents mainly reside with their sons rather than daughters.[17] The interpretation is that when they coreside, a fraction of children's income is transferred to parents. In the section on model calibration, I will map the monetary transfer to intergenerational coresidence and show that my specification can replicate the pattern of coresidence very well.

The total consumption at old age is the savings from middle age and the transfers from children, i.e.,

$$c_3 = a(1 + r) + q_m t_m(q_m, q_f) I_{2m} + q_f t_f(q_m, q_f) I_{2f}, \tag{10}$$

[16]In the model, old-age support is assumed to depend only on the number of sons and daughters. In reality, it may also depend on other sources of parental income. For example, Cheng et al. (2018) find that the expansion of the New Rural Pension Scheme in China increased the likelihood that parents live independently by raising their income, and Mukherjee (2020) shows that children's caregiving is decreasing in parental social security benefits by exploiting the social security notch in the US.

[17]The data from the Longitudinal Ageing Study in India in 2017-2018 suggest that less than 5% of those aged above 60 years received a monetary transfer from children, but more than 60% of them were residing with their children. Coresidence with sons is not unique to India. Instead, it is present in many developing countries and has been argued as an important reason why parents want more sons than daughters (e.g., Das Gupta et al., 2003; Chung and Das Gupta, 2007; Ebenstein, 2021).

where I_{2m} is the household income net of dowry and savings of each son, and I_{2f} is its counterpart of each daughter.

Single Households

As will be clear later, there are some single male households in the economy. Consider a single male with the human capital level h_m. When he is young and middle-aged, he supplies one unit of time to the market and earns labor income h_m. In addition, he receives a lump-sum transfer T, which is the redistribution of the discrimination costs. When young, the income is fully consumed. Singles do not consume any home production. When middle-aged, the male saves a for the old age. A fraction of the remaining income is transferred to parents, as specified by Equation (8). When old, the male consumes his savings, which amounts to $a(1+r)$.

3.2 Problems of Married Households

Household problems are solved using backward induction. I start by computing the value of being old and then look at the decisions on savings and female labor supply of middle-aged adults. After that, I solve for the decisions on education and female labor supply of young adults for a given number of sons, daughters, and children born with sex selection. Finally, I look at the sequential fertility and sex selection decisions.

Old households make no economic decisions. They simply consume their savings and transfers from their children, which is determined by the current state variables $\mathbf{x}_3 = \{a, q_m, q_f, \mathbf{x}_2^{k,m}, \mathbf{x}_2^{k,f}\}$, where $\mathbf{x}_2^{k,m}$ and $\mathbf{x}_2^{k,f}$ are the state of their male and female children, who are middle-aged. Therefore, the value of being old is

$$V_3(\mathbf{x}_3) = \log[(1+r)a + q_m t_m(q_m, q_f) I(\mathbf{x}_2^{k,m}) + q_f t_f(q_m, q_f) I(\mathbf{x}_2^{k,f})], \tag{11}$$

where $I(.)$ is given by Equation (6) for the male and female children, who are now middle-aged.

For a middle-aged household, the state variables is $\mathbf{x}_2 = \{h_m, h_f, q_m, q_f, \mathbf{x}_3^{p,m}, \mathbf{x}_3^{p,f}\}$, where $\mathbf{x}_3^{p,m}$ and $\mathbf{x}_3^{p,f}$ are the state of the parents of the husband and the wife, which include $\{q_m^{p,m}, q_f^{p,m}\}$ and $\{q_m^{p,f}, q_f^{p,f}\}$, i.e., the number of male and female children in the natal family of the husband and the wife, respectively.

Given $\mathbf{x}_2$, households choose $\{l_2, n_2, c_2, a\}$ to maximize the value of being middle-aged,

$$V_2(\mathbf{x}_2) = \max_{\{l_2,n_2,c_2,a\}} \log(c_2) + \nu_2 \log(n_2) + \beta \mathbb{E}[V_3(\mathbf{x}_3)], \tag{12}$$

subject to

$$c_2 = [1 - t_m(q_m^{p,m}, q_f^{p,m}) - t_f(q_m^{p,f}, q_f^{p,f})][h_m(1 - q_f\delta) + \lambda h_f l_2 + T - a],$$

and

$$n_2 = 1 - l_2.$$

In the objective function, ν_2 governs the relative importance of consumption of non-market goods relative to consumption of market goods in middle age. Note that when households predict $\mathbf{x}_3$, they predict the middle-aged status of their kids, since $\mathbf{x}_2^{k,m}$ and $\mathbf{x}_2^{k,f}$ are part of $\mathbf{x}_3$.

Now, look at the problem of young adults. First suppose that the number of sons, daughters, and children born with sex selection, $\{q_m, q_f, q_s\}$, are determined and that parents will not have more children. Then parents only make decisions on education investment and female labor supply. The state variables for the couple is given by $\mathbf{x}_1 = \{h_m, h_f, q_m, q_f, q_s, \mathbf{x}_2^{p,f}\}$, where $\mathbf{x}_2^{p,f}$ determines the dowry payments by the wife's parents to the young couple. The couple chooses $\{l_1, n_1, e_m, e_f, c_1\}$ to solve the following problem,

$$V_1(\mathbf{x}_1) = \max_{\{l_1,n_1,e_m,e_f,c_1\}} \log(c_1) + \nu_1 \log(n_1) + U_a(q_m, q_f, e_m, e_f) - q_s\xi + \beta \mathbb{E}[V_2(\mathbf{x}_2)], \tag{13}$$

subject to

$$c_1 = h_m + \lambda h_f l_1 + d(\mathbf{x}_2^{p,f}) + T - q_m e_m - q_f e_f - q_s \pi,$$

$$n_1 = 1 - l_1 - (q_m + q_f)\tau,$$

and

$$Prob(h'_g = h_{g,h} | e_g) = 1 - \exp(-\eta e_g), g \in \{m, f\}.$$

In the objective function, ν_1 governs the relative importance of consumption of non-market goods relative to consumption of market goods when young. ν_1 can be different from ν_2, implying that people may have different tastes for domestic goods in different ages. Note that, in the budget constraint for market goods, $d(\mathbf{x}_2^{p,f})$ is the dowry payments from wife's parents. The last constraint is the production technology of human capital.

Now we know the value of ending up with q_m sons and q_f daughters, among which q_s children were born with the sex selection technology. Note that the other variables in the state vector $\mathbf{x}_1$ except $\{q_m, q_f, q_s\}$ cannot be chosen by the household. For a better illustration, in the following I will use $V_1(q_m, q_f, q_s)$ rather than $V_1(\mathbf{x}_1)$ to denote the value of ending up with $\{q_m, q_f, q_s\}$.

Now we can look at parental decisions on fertility and sex selection. These decisions are made sequentially conditional on the current number of sons (q_m), daughters (q_f) and children born with sex selection (q_s). Denote the value of having $\{q_m, q_f, q_s\}$ by $V(q_m, q_f, q_s)$. The problem can be solved using backward induction.

Denote the maximum possible number of children by $\bar{q}$. First consider a couple with $\bar{q}$ children, i.e., $q_m + q_f = \bar{q}$. Since they cannot have more children, the value of $\{q_m, q_f, q_s\}$ is $V_1(q_m, q_f, q_s)$. That is,

$$V(q_m, q_f, q_s) = V_1(q_m, q_f, q_s), \quad \text{if} \quad q_m + q_f = \bar{q}. \tag{14}$$

Next, consider a couple with $\{q_m, q_f, q_s\}$, where $q_m + q_f = \bar{q} - 1$. The value of *not* having another birth is

$$V_{nb}(q_m, q_f, q_s) = V_1(q_m, q_f, q_s). \tag{15}$$

The value of *having* another birth *without* sex selection is

$$V_b(q_m, q_f, q_s) = p_m V_1(q_m + 1, q_f, q_s) + (1 - p_m) V_1(q_m, q_f + 1, q_s). \tag{16}$$

The value of *having* another birth with a *male* child being selected is

$$V_{bm}(q_m, q_f, q_s) = V_1(q_m + 1, q_f, q_s + 1). \tag{17}$$

The value of *having* another birth with a *female* child being selected is

$$V_{bf}(q_m, q_f, q_s) = V_1(q_m, q_f + 1, q_s + 1). \tag{18}$$

Parents choose whether to have another birth ($b \in \{0, 1\}$) and whether to select a male birth ($s_m \in \{0, 1\}$) or a female birth ($s_f \in \{0, 1\}$) if they will have another one. Therefore, the value of $\{q_m, q_f, q_s\}$ if $q_m + q_f = \bar{q} - 1$ is

$$\begin{aligned} V(q_m, q_f, q_s) = \max_{\substack{b, s_m, s_f \in \{0,1\}, \\ s_m + s_f \in \{0,1\}}} & (1 - b) V_{nb}(q_m, q_f, q_s) \\ & + b[(1 - s_m - s_f) V_b(q_m, q_f, q_s) + s_m V_{bm}(q_m, q_f, q_s) + s_f V_{bf}(q_m, q_f, q_s)]. \end{aligned} \tag{19}$$

Denote the fertility and sex selection choices by $b(q_m, q_f, q_s)$, $s_m(q_m, q_f, q_s)$ and $s_f(q_m, q_f, q_s)$. Now we have the policy functions and the value function if the number of children is $\bar{q} - 1$. Going backward, we can solve for the fertility and sex selection decisions if the number of children is $\bar{q} - 2$, $\bar{q} - 3$, ..., 0 in the same way.

3.3 Marriage Market

Young unmarried adults meet in a marriage market. Denote the measure of young adults of sex g with human capital h by Q_g^h ($g \in \{m, f\}$). Since $p_m > 0.5$ and parents tend to select more male than female children, we have

$$Q_f^{h_l} + Q_f^{h_h} < Q_m^{h_l} + Q_m^{h_h}. \tag{20}$$

To replicate assortative mating in the data, the marriage market is assumed to clear in three stages, following Fernández and Rogerson (2001), Fernandez et al. (2005) and Caucutt et al. (2021). First, a fraction of males with the lowest education are excluded from the market randomly such that the remaining males and females have the same measure. Denote the measure of remaining young adults of sex g with human capital h by $\tilde{Q}_g^h$, then we have

$$\tilde{Q}_f^{h_l} + \tilde{Q}_f^{h_h} = \tilde{Q}_m^{h_l} + \tilde{Q}_m^{h_h}. \tag{21}$$

Next, males and females with the same education meet, and some of them get matched. The measure of matched males/females is $\psi min\{\tilde{Q}_m^h, \tilde{Q}_f^h\}$. The parameter ψ governs the degree of assortative mating. If $\psi = 0$, we have random matching, while if $\psi = 1$, we have perfect sorting. At the last stage, those who were not matched assortatively are matched randomly.

Finally, we arrive at a distribution of households. In the analysis, I focus on the stationary equilibrium. In a stationary equilibrium, given the state variables, young households make decisions on fertility, sex selection, education investment, and female labor supply to maximize their utility; middle-aged households make decisions on female labor supply and savings. The marriage market clears, and the distribution of households and household decisions are constant over time.[18]

[18] To find a stationary equilibrium, a computational fixed point problem is solved. First, an initial guess is made about the lump-sum transfer from the discrimination cost, the distribution of households, the distribution of sons' and daughters' household types given one's household type, and transfers from

4 Quantitative Analysis

In this section, I discuss the calibration procedures and the model fit. One period in the model corresponds to 20 years, and thus, an individual in the model lives 80 years. Parents are assumed to have at most $\bar{q} = 6$ children. Low human capital in the model corresponds to at most primary education in the data, while high human capital individuals are those with higher levels of education. In the DHS data in 2015-2016, for individuals aged 21-54, those with low education receive about 1.6 years of schooling on average, while those with high education receive about 10.5 years of schooling. 34.3% of males and 46.7% of females have low human capital, while the rest have high human capital.

4.1 Calibration

I first set some parameters to their data counterparts or borrow them from the literature. Given the natural ratio of male to female births around 1.05, the probability of a male birth in the absence of sex selection, p_m, is set to 0.5122. The annual interest rate is set to 5.83%, which is the average of the real interest rate in India in 1978-2020 from the World Bank.[19]

To compute human capital, or wages rates, associated with two education levels $\{h_l, h_h\}$, I divide the total earnings (wage earnings, farm income, and business income) by the working hours to get the hourly earnings for each male aged 21-60 using data from the Indian HDS in 2011-2012. I restrict the sample to those working more than 200 hours and run a Mincerian regression where the dependent variable is log(hourly earnings) and the key independent variable is a dummy variable indicating whether the individual completed secondary education. Meanwhile, I control for individuals' age, age squared, marital status, state, religion, and caste. The results suggest that the wage rate of men with high education is about 1.586 times that of men with low education. Therefore, I

sons and daughters conditional on their household types. Given the initial guess, household problems are solved, and the marriage market is cleared. Then the guessed variables are updated. The process continues until all the variables converge. Although it is not possible to establish the uniqueness of the fixed point theoretically, the solution algorithm converges to the same point from several different initial guesses.

[19]See https://data.worldbank.org/indicator/FR.INR.RINR?locations=IN.

normalize h_l to 1 and set h_h to 1.586.[20]

The labor market discrimination parameter λ is set to 0.67, following Deshpande et al. (2018). They use the Blinder-Oaxaca method to decompose the gender wage gap into an "explained part" (due to gender differences in wage-earning characteristics) and an "unexplained component" (due to gender differences in the labor market returns to characteristics). Their results suggest that the unexplained gender gap in log(wage) is 0.37 in 1999-2000 and 0.43 in 2009-2010, implying that women earn 65-69% of men's wage, *ceteris paribus*. Therefore, I set λ to 0.67.[21]

The degree of assortative mating (the fraction of individuals who meet someone with the same education level as themselves), ψ, is set to 0.557. I use the sample of married women aged 21-49 and their spouses in the DHS data in 2015-2016.[22] In the data, 27% of all marriages are between two low-human capital spouses, while 46% are formed by two high-human capital ones. The share of mixed couples is 19% (a low-human capital wife with a high-human capital husband) and 7% (a high-human capital wife and a low-human capital husband). The parameter ψ is set to replicate these shares.

The level of dowry as a fraction of paternal income δ is computed based on the Rural Economic and Demographic Survey (REDS) in 1999. I look at dowry payments in the last 5 years before the survey and divide them by the household income in the last year, which gives a median dowry-to-income ratio of 0.5. Since women's share of the household income is about 16.7% in the HDS data in 2011-2012, this implies that the ratio of dowry to the father's income is about 0.6. Given that the model period is 20 years, δ is set to $0.6/20 = 0.03$.

The monetary cost of sex selection π is set to 0.005. In the model, it is assumed that parents can directly select a male/female birth. In reality, parents usually combine prenatal sex detection and abortion to pursue a male/female birth. Parents first diagnose the sex of the fetus, and if the sex is unwanted, they abort the fetus. They repeat this

[20] This number is consistent with the results from Agrawal and Agrawal (2019), who estimate returns to education measured by years of schooling in India using the same data.

[21] In principle, the unexplained part should not be fully interpreted as discrimination, since there may also be gender differences in unobserved wage-earning characteristics. However, Deshpande et al. (2018) take into account a comprehensive set of characteristics, including age, caste, marital status, education, rural/urban residence, regions, private/public sector, union membership, temporary/permanent job, occupation, and industry. Therefore, the unexplained gender wage gap is plausibly discriminatory.

[22] Women in the Indian DHS sample are 15-49 years old. I drop those aged below 21 years since a large fraction of them (79.8%) had never married at the time of the survey.

process until they conceive a fetus of wanted sex. Therefore, the monetary cost of sex selection consists of the cost of prenatal sex detection and the cost of abortions. To have a male birth, they need to conduct on average $1/p_m = 1.95$ tests and $1/p_m - 1 = 0.95$ abortions. To have a girl, they need to conduct $1/(1-p_m) = 2.05$ tests and $1/(1-p_m)-1 = 1.05$ abortions. For simplicity, assume that they need to conduct two tests and one abortion regardless of the targeted sex. The cost of sex detection via ultrasound is 500-1000 rupees per test (about 10-20 dollars at the current rate of exchange) in the 1990s (Arnold et al., 2002).[23] I take the middle value of 750 rupees. The cost of an abortion in the second trimester is 1661 rupees (about 35 dollars) on average in 1996–1998 (Ganatra and Hirve, 2002).[24] Therefore, the expected cost of sex selection is $750\times2+1661\times1 = 3111$ rupees. Since the male income with low education is normalized to 1 in the model, we need to normalize the monetary cost of sex selection as well. In the Indian HDS data in 2011-2012, the median annual income of males with low education is 30854 rupees (about 617 dollars).[25] Since the model period is 20 years, the cost of each sex selection by the parents is about $3111/(30854 \times 20) = 0.005$ of their one model-period income.

Finally, I calibrate the parameters governing old-age supports to parents, $\{t_{m1}, \kappa_m, t_{f1}, \kappa_f\}$. Recall that in India, parents primarily receive old-age support from their children by residing with them. As a result, I chose the functional forms for old-age transfers to capture the intergenerational coresidence patterns in the data. Suppose that an exogenous fraction ϕ of children's income flows to parents while they coreside. Furthermore, let the fraction of time that parents spend with each son and daughter be denoted by R_m and R_f, respectively. Then, the fraction of income that children transfer to their parents is $t_g(q_m, q_f) = \phi R_g(q_m, q_f)$ for $g \in \{m, f\}$. Therefore, the fraction of time spent with children can be written as

$$R_m(q_m, q_f) = t_m(q_m, q_f)/\phi = \begin{cases} R_{m1} q_m^{-\kappa_m}, & \text{if } q_m > 0, \\ 0, & \text{if } q_m = 0, \end{cases} \tag{22}$$

[23] Another method to detect the sex of fetuses is amniocentesis, which is invasive and more expensive (Arnold et al., 2002). Therefore, I consider ultrasound only.

[24] The sex of a fetus cannot be well detected in the first trimester.

[25] I use the median rather than the mean since it is less sensitive to outliers.

and,

$$R_f(q_m, q_f) = t_f(q_m, q_f)/\phi = \begin{cases} 0, & \text{if } q_m > 0, \\ R_{f1} q_f^{-\kappa_f}, & \text{if } q_m = 0, \end{cases} \tag{23}$$

where, $R_{m1} = t_{m1}/\phi$ is the time of coresidence if parents have only a son, and $R_{f1} = t_{f1}/\phi$ is the time of coresidence if parents have only a daughter. The parameters κ_m and κ_f capture how the time changes with the number of children, as in Equations (8) and (9).

I use data from the Longitudinal Ageing Study in India (LASI) in 2017-2018 to estimate $\{R_{m1}, \kappa_m, R_{f1}, \kappa_f\}$, and given an estimate for ϕ to recover $\{t_{m1}, \kappa_m, t_{f1}, \kappa_f\}$. In particular, I estimate the total time spent with children (time spent with each child times the number of children) R as

$$R = \begin{cases} R_{m1} q_m^{1-\kappa_m}, & \text{if } q_m > 0, \\ R_{f1} q_f^{1-\kappa_f}, & \text{if } q_m = 0. \end{cases} . \tag{24}$$

Hence, as long as parents have any sons, they reside with them, and the amount of time they spend with them increases in the number of sons they have. If they do not have any sons, then they reside with their daughters, with the amount of time they spend again increasing in the number of daughters.

In the data, the time of intergenerational coresidence is not available, but one can observe whether parents were residing with their children (a son or daughter) at the time of the survey, which I use as a proxy for the time of coresidence. I run two non-linear least square regressions, one for parents with sons and the other for parents with only daughters. The dependent variable is whether a parent is residing with his/her children at the time of the survey, and the independent variable is the number of sons/daughters. I restrict the sample to parents aged 61 years and above. The results show that $R_{m1} = 0.713$, $\kappa_m = 0.972$, $R_{f1} = 0.409$, and $\kappa_f = 0.925$.

Table 3 shows the share of parents residing with a child by the number of sons and daughters in the data and the corresponding fitted values given by the non-linear least square regressions. Equation (24) captures the coresidence pattern in the data very well. First, the probability of residing with children is much lower for parents with only daughters than for parents with sons. Second, once parents have sons, the probability of residing with children increases very slowly in the number of sons.[26]

There is no direct evidence on ϕ. To obtain t_{m1} and t_{f1}, I assume that ϕ is 0.29, which is consistent with the square root equivalence scale for households with four adults (Eckstein et al., 2019). This leads to $t_{m1} = R_{m1}\phi = 0.207$ and $t_{f1} = R_{f1}\phi = 0.119$. This means that an only son will transfer 20.7% of his household income to his parents, while an only daughter will transfer 11.9% of her household income to her parents. The parameters chosen outside the model are summarized in Table 4.

Table 3: Share of Parents Coresiding with Children: Data and Model

	0 sons	1 son	2 sons	3 sons	4 sons	5 sons
0 daughters	0	0.75	0.74	0.73	0.80	0.66
	(0)	(0.71)	(0.73)	(0.73)	(0.74)	(0.75)
1 daughter	0.39	0.69	0.71	0.72	0.72	0.77
	(0.41)	(0.71)	(0.73)	(0.73)	(0.74)	(0.75)
2 daughters	0.43	0.71	0.71	0.79	0.69	0.87
	(0.43)	(0.71)	(0.73)	(0.73)	(0.74)	(0.75)
3 daughters	0.45	0.77	0.74	0.83	0.85	0.77
	(0.44)	(0.71)	(0.73)	(0.73)	(0.74)	(0.75)
4 daughters	0.58	0.72	0.82	0.76	0.91	0.78
	(0.45)	(0.71)	(0.73)	(0.73)	(0.74)	(0.75)
5 daughters	0.46	0.85	0.82	0.82	0.79	0.79
	(0.47)	(0.71)	(0.73)	(0.73)	(0.74)	(0.75)

Notes. Numbers in parentheses are the fitted values of the non-linear least square regressions.

[26]For this reason, parents may want at least one or two sons but not necessarily more, which is the observation of Jayachandran (2017).

Table 4: Parameters Chosen outside the Model

	Description	Value	Source
$\bar{q}$	Maximum number of children	6	
p_m	Natural probability of a male birth	0.5122	
r	Annual interest rate	5.83%	World Bank
h_l, h_h	Human capital	1, 1.586	Indian HDS 2011-2012
λ	Labor market discrimination	0.67	Deshpande et al. (2018)
ψ	Assortative mating	0.557	Indian DHS 2015-2016
δ	Dowry as a fraction of parental income	0.03	REDS 1999
π	Monetary cost of sex selection	0.005	Arnold et al. (2002), Ganatra and Hirve (2002)
t_{m1}, κ_m	Transfer from sons (if $>$ 0 son)	0.207, 0.972	LASI 2017-2018
t_{f1}, κ_f	Transfer from daughters (if 0 son)	0.119, 0.925	LASI 2017-2018

4.2 Estimation and Model Fit

We are left with 13 undetermined parameters: 11 parameters governing household preferences $\{\beta, \nu_1, \nu_2, \mu_\alpha, \sigma_\alpha, \omega, \epsilon, \eta, e_0, \mu_\xi, \sigma_\xi\}$, the efficiency of human capital production (z), and the fixed time cost of each child (τ). These parameters are estimated within the model to match the following 48 data moments:

(i) Household savings rate when middle-aged. Source: Indian HDS 2011-2012. [1 target]

(ii) Women's working hours when young and middle-aged as a fraction of their time endowment. Source: Indian HDS 2011-2012. [2 targets]

(iii) Parity progression ratios conditional on the sex composition of existing children, as shown by Fact 1 in Section 2. Source: Indian DHS 2015-2016. [20 targets]

(iv) Fertility difference between women with and without secondary education. Source: Indian DHS 2015-2016. [1 target]

(v) Share of married men and women with secondary education. Source: Indian DHS 2015-2016. [2 targets]

(vi) Education expenditure per son and per daughter as a fraction of household income. Source: Indian HDS 2011-2012. [2 targets]

(vii) Probability that the next child is a son conditional on the sex composition of existing children, as shown by Fact 2 in Section 2. Source: Indian DHS 2015-2016. [20 targets]

One parameter may affect multiple moments and one moment may be affected by multiple parameters, so it can be difficult to determine which moment pins down which parameter. In Appendix B, a systematic analysis is conducted to examine how the moments affect the parameters using the method proposed by Andrews et al. (2017). The following discussion mainly reflects this more formal analysis. Intuitively, the discount factor β determines the savings rate. ν_1 and ν_2 affect women's labor supply, but ν_1 may also affect fertility decisions since each birth requires a time commitment of the mother. The distribution of weight of children $\{\mu_\alpha, \sigma_\alpha\}$ affects the parity progression ratios, while the distribution of disutility from sex selection $\{\mu_\xi, \sigma_\xi\}$ affects the sex selection decisions.

The relative importance of children (η), public education expenditure (e_0), and fixed time cost of each birth (τ) influence women's education fertility gradient and education expenditure per son and per daughter by altering parental quantity-quality trade-off. Moreover, all decisions on fertility, sex selection, and education are affected by the relative importance of daughters (ω) and the elasticity of substitution between sons and daughters (ϵ). Finally, the efficiency of human capital production (z) determines the link from education expenditure for boys and girls to the distribution of men's and women's education.

The estimated parameters are shown in Table 5. The elasticity of substitution between sons and daughters (ϵ) is larger than one, suggesting that sons and daughters are imperfect substitutes. Therefore, parents tend to have another birth if the sex composition of their children is very imbalanced, i.e., too few sons or too few daughters. The weight of daughters relative to sons (ω) is smaller than one, suggesting that parents have an intrinsic preference for sons. As a result, parents are more likely to have another birth if sons make up a small proportion of existing children than if daughters make up a small proportion. The values of ϵ and ω also imply that parents may select a male birth after they have some daughters but no sons. The value of τ implies that each birth takes about 11% of the mother's time.

Table 5: Estimated Parameters

	Description	Value	Target
β	Annual discount factor	0.9765	Moments (i)
ν_1	Weight of non-market goods when young	0.2274	Moments (ii)
ν_2	Weight of non-market goods when middle-aged	0.5571	Moments (ii)
μ_α, σ_α	Distribution of weight of children	-1.1239, 0.5402	Moments (iii)-(vii)
ω	Relative importance of daughters	0.9237	Moments (iii)-(vii)
ϵ	Substitution btw sons and daughters	19.93	Moments (iii)-(vii)
μ_ξ, σ_ξ	Distribution of disutility from sex selection	-3.5493, 0.4464	Moments (iii)-(vii)
η	Importance of education investment	0.5795	Moments (iii)-(vii)
e_0	Public education expenditure	0.0311	Moments (iii)-(vii)
z	Efficiency of human capital production	12.13	Moments (iii)-(vii)
τ	Fixed time cost of each child	0.1066	Moments (iii)-(vii)

Notes. 48 moments are used to estimate 13 parameters.

Moments (i), (ii), (iv), (v), and (vi) in the data and model are shown in Table 6. Moments (iii) and (vii) are shown in Figures 4 and 5, respectively. The model can replicate the data well.[27] In particular, it can replicate the son-biased fertility stopping rule and

[27]The model predicts a SRB of 1.139, which is smaller than the data moment (1.184). In fact, women

the son-biased sex selections.

Table 6: Data and Model: Selected Moments

Variable	Data	Model
Targeted moments		
Savings rate when middle-aged	0.367	0.361
Female labor supply when young	0.235	0.250
Female labor supply when middle-aged	0.264	0.270
Fertility difference across women's education	0.93	0.81
Share of married men with high education (%)	65.7	73.6
Share of married women with high education (%)	53.3	49.2
Education expenditure per son	0.0505	0.0570
Education expenditure per daughter	0.0426	0.0379
Aggregate moments		
Fertility	3.04	3.07
Sex ratio at birth	1.184	1.139

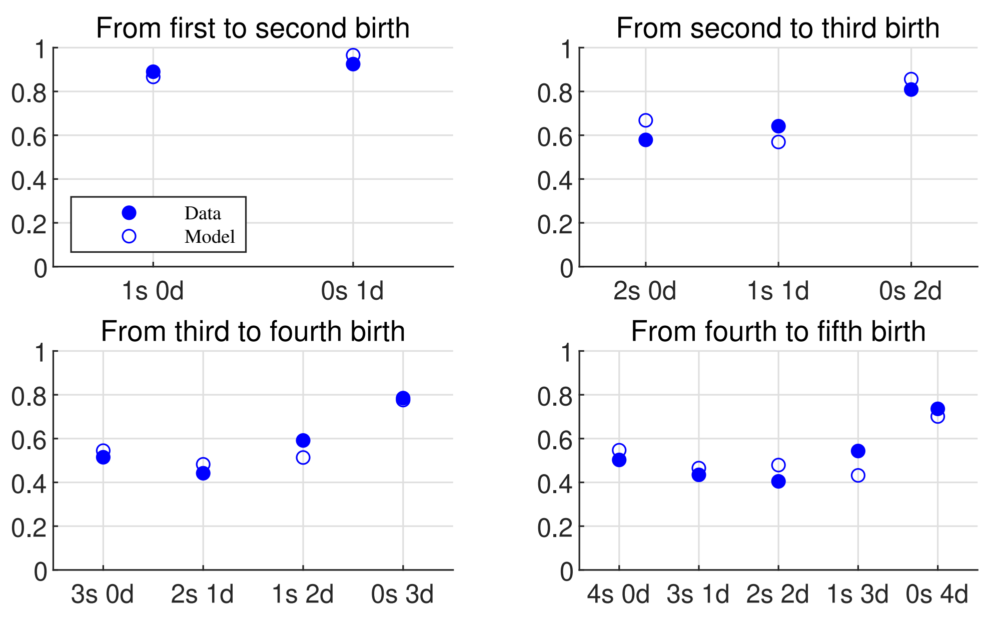


Figure 4. Parity Progression Ratios Conditional on the Sex Composition of Previous Children: Data and Model.

tend to underreport female children that did not survive, so the moment in the data is probably overestimated (Rosenblum, 2013). Based on data from the 2002–2004 Reproductive and Child Health Survey, Rosenblum (2013) suggests that the overestimation of the SRB for the first births amounts to 0.023, although the overestimation for later births is smaller.

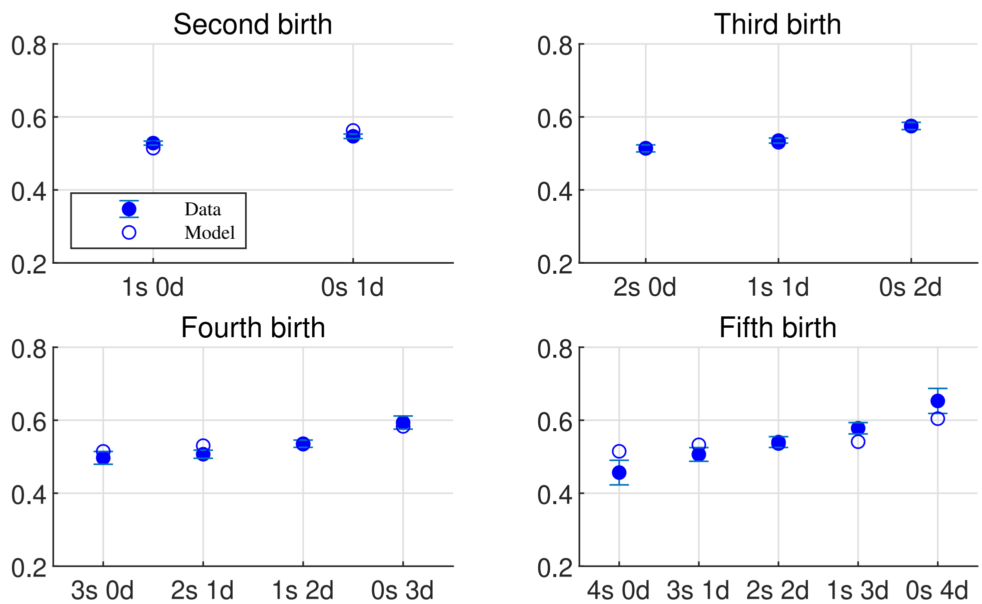


Figure 5. Probability of a Son Conditional on the Sex Composition of Previous Children: Data and Model.

4.3 Non-targeted Moments

This subsection shows the model fit of some non-targeted moments, i.e., the ones not used in estimation. Figures 6 and 7 show the PPR and the probability that the next child is a son conditional on the sex composition of existing children for women with different education. Table 7 presents fertility, the sex ratio at birth, and education expenditure per son and per daughter for women with different education. Table 8 reports the distribution of married men, married women, and couples based on their education. The results suggest that the model can match most of the moments very well, lending more credibility to the quantitative predictions of the model. One exception is that the model predicts that women with high education are more likely to select a male birth than women with low education, while the data suggest no significant difference.[28,29]

[28]The relationship between education and sex selection is not straightforward. It depends on both preferences and constraints of parents (Bongaarts, 2013). On the one hand, a woman's education may alter her perception of "feminine worth," thus reducing her preference for sons (Pande and Astone, 2007). On the other hand, sex selection is more affordable for more educated women. Meanwhile, they have more incentives to select a male birth because they have fewer children and are more likely to end up with no son if they do not select their children's sex (Jayachandran, 2017). The model captures the latter mechanism but leaves out the former.

[29]It should be noted that the SRB in the data is probably overestimated since women tend to underreport female children that did not survive (Rosenblum, 2013). The bias is expected to be larger for women with low education since women's education is negatively related to the gender gap in child mortality (Bourne and Walker, 1991).

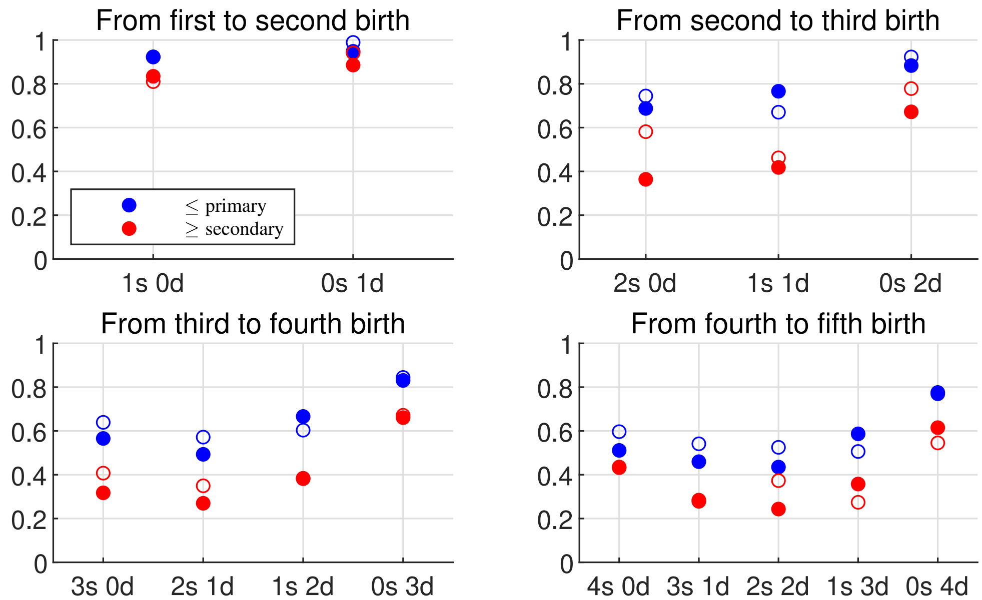


Figure 6. Parity Progression Ratios Conditional on the Sex Composition of Previous Children by Women's Education: Data and Model.

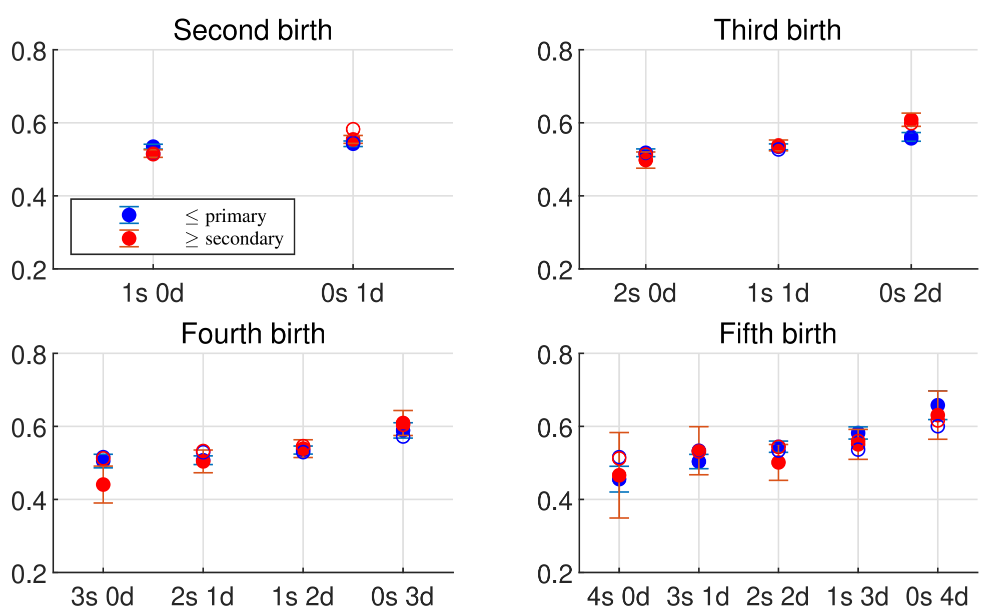


Figure 7. Probability of a Son Conditional on the Sex Composition of Previous Children by Women's Education: Data and Model.

Table 7: Data and Model: Selected Non-targeted Moments

Variable	Data	Model
Fertility (women with low education, henceforth low)	3.38	3.47
Fertility (women with high education, henceforth high)	2.45	2.66
Sex ratio at birth (low)	1.185	1.118
Sex ratio at birth (high)	1.180	1.168
Education expenditure per son (low)	0.0439	0.0478
Education expenditure per son (high)	0.0619	0.0665
Education expenditure per daughter (low)	0.0336	0.0285
Education expenditure per daughter (high)	0.0574	0.0475

Table 8: Distribution of Married Men, Women, and Couples

		Men	
		Primary or lower	Secondary or higher
Women	Primary or lower	0.27 (0.22)	0.19 (0.29)
	Secondary or higher	0.07 (0.04)	0.46 (0.45)

Notes. Model predictions in parentheses.

4.4 External Validation

In this section, the model is used to simulate sex selection and fertility responses to a conditional cash transfer program named *Dhanlakshmi*, which was implemented in seven Indian states in 2008-2013. The model predictions are compared with the empirical findings from Biswas et al. (2021). The purpose of this exercise is to show that the model delivers credible predictions of household behaviors.

In 2008, *Dhanlakshmi* was introduced in 11 blocks (the next administrative level below the district) across seven states: Andhra Pradesh, Bihar, Chhattisgarh, Jharkhand, Odisha, Punjab, and Uttar Pradesh. The program aimed to normalize the sex ratio, increase the education and health investment of girls, and delay their marriage. All girls born on or after November 19, 2008 could be enrolled in this program regardless of household income or the number of siblings. Parents of a girl could receive 5000 rupees (about 115 dollars at the current rate of exchange) by showing her birth registration. They could receive 1250 rupees (about 29 dollars) in addition if the girl was fully immunized. Furthermore, parents could receive a total of 3500 rupees (about 81 dollars) if the girl completed primary education (Grade 5) and an extra transfer of 3750 rupees (about 86 dollars) if she completed secondary education (Grade 8). Finally, if the girl remained unmarried until 18 years, she could receive an insurance maturity cover of 100,000 rupees (about 2302 dollars). The program was stopped in 2013 but the financial commitment continued to be honored.

Dhanlakshmi is an ideal program as a validation exercise for the current model. First, the participation cost was very low. Unlike other programs that restricted the number of children of beneficiary parents and required them to get sterilized (e.g., *Rajalakshmi* in Rajasthan and *Devi Rupak* in Haryana) or those targeted at specific income groups (e.g., *Apni Beti Apna Dhan* in Haryana for those below the poverty line), *Dhanlakshmi* did not impose any requirement on household income or the number of siblings of enrolled girls. The only eligibility criterion was that the family resided in areas where the program was implemented. Second, the transfers were sizable. To put the benefits in perspective, recall that the median annual income was 30854 (about 617 dollars) rupees for males without secondary education in the HDS data in 2011-2012. With sizable benefits, this program was likely to cause observable changes in fertility and sex selections both in the data and in the model.

Biswas et al. (2021) examine how this program affected sex selection (measured by the probability that a birth is a female) and fertility (measured by whether a woman has a birth in a given year) using a difference-in-difference approach. More specifically, they construct a woman-year panel for the years 2004-2007 (before the policy) and 2009-2012 (after the policy). They focus on the state of Punjab and compare women in Fatehgarh Sahib (the treatment block in Punjab) and women in the other blocks in this state (control blocks). They find that the program increased the probability that a birth is a female by 4.5-5.5 pp. In addition, it increased the probability for a woman to have a birth in a given year from 0.287 by 0.0051-0.0085. The DHS data in 2015-2016 suggest that an average woman in Punjab would have about 2.65 births in her life. This implies that this program would increase the fertility rate by about 0.047-0.078.[30].

To simulate the program in the model, I focus on the transfers for female births and

[30] To get the increase in the fertility rate, we need to multiple the percentage increase in the annual birth risk by the life-time fertility rate, i.e., $0.051/0.287 \times 2.65 = 0.047$ and $0.085/0.287 \times 2.65 = 0.078$.

education.[31] Since the male income with low education is normalized to 1 in the model, we need to normalize the transfers as well. As mentioned above, in the HDS 2011-2012 survey, the median annual income of males with low education is 30854 rupees. Since the model period is 20 years, the transfer for a female birth is about 5000/30854/20 = 0.0081 in the model. Considering that everyone can achieve low education in the model, which corresponds to about 1.6 years of education in the data, I assume that in the model parents cannot receive any transfer if their daughter completes low education and that they can receive a total of 7250 rupees if their daughter completes high education. This transfer amounts to 7250/30854/20 = 0.0117 in the model. Finally, it is assumed that the transfer for female births is received in the young period, while the transfer for female education is received in the middle age, when daughters' education achievement is realized.

It should be noted that the sex ratio at birth and fertility rate in Punjab are different from those in an average Indian state. In particular, the sex ratio is more biased in Punjab. To make the simulation results comparable to the empirical findings of Biswas et al. (2021), the model is recalibrated to the Punjab data. The details of the calibration are explained in Appendix C.

The *Dhanlakshmi* program is first simulated in a partial equilibrium (PE) where the distribution of households, parental expectations about children's marriages and transfers, etc., remain unchanged. Next, it is simulated in a stationary equilibrium (GE) where the transfers are assumed to be financed by a proportional labor income tax on currently working men.

The results are displayed in Table 9. First, the model predicts that the probability of a female birth would be increased by 3.8 pp in the PE and 3.6 pp in the GE, which are close to the estimate of Biswas et al. (2021), i.e., 4.5-5.5 pp. Second, the model predicts that fertility would be increased by 0.173 in the PE and 0.131 in the GE, which are slightly larger than the estimate of Biswas et al. (2021), i.e., 0.047-0.078. Overall, the model predictions are close to the empirical findings of Biswas et al. (2021), suggesting that the model can deliver reliable predictions in counterfactual analyses.

[31] Certainly, the transfers for immunization and delay of marriage may also affect fertility and sex selection. However, it is difficult to incorporate them in the model, which abstracts from health investment and age of marriage.

Table 9: Effects of *Dhanlakshmi*

	Biswas et al. (2021)	Model (PE)	Model (GE)
Prob of a female birth before the policy (%)	45.6	43.0	43.0
Increase in the prob of a female birth (pp)	4.5-5.5	3.8	3.6
Annual fertility risk before the policy	0.287		
Increase in annual fertility risk	0.0051-0.0085		
Fertility before the policy	2.65[a]	2.53	2.53
Increase in fertility	0.047-0.078	0.173	0.131

Notes. *a*: Author's estimate based on the data from the Indian DHS in 2015-2016.

4.5 Sex Selection Technology

What is the impact of sex selection technology in the benchmark economy? To answer this question, the monetary cost of sex selection in the model is raised to infinity so that nobody can afford it.[32] The simulation results for the key variables are reported in Table 10.

Table 10: Impacts of Sex Selection Technology

	Benchmark	No sex selection
Fertility	3.07	3.09
Sex ratio at birth	1.139	1.050
Education expenditure per son	0.0570	0.0570
Education expenditure per daughter	0.0379	0.0376
Men with high education (%)	64.6	63.6
Women with high education (%)	49.2	48.2
Gender gap in education	15.4	15.4
Female labor (young)	0.250	0.254
Female labor (middle-aged)	0.270	0.279

The sex ratio at birth is reduced to the natural level by construction. However, its effects on other variables are relatively small. There is a decrease in education investment in children. In particular, education expenditure per daughter is reduced by 1%, and

[32]This experiment is a complete ban on sex selection and mimics the Pre-Conception and Pre-Natal Diagnostic Techniques Act of 1994, which aimed at banning prenatal sex detection and preventing sex-selective abortions. Although this policy was difficult to implement in practice (Das Gupta, 2019), Nandi and Deolalikar (2013) find that it did reduce the ratio of male to female births. In this counterfactual experiment, it is assumed that the sex selection ban can be effectively implemented.

the share of women with high education is reduced by 1.0 pp. This reflects a trade-off between the quantity and the quality of daughters. The findings about sex ratio and expenditure for girls are consistent with the empirical findings in the literature. Nandi and Deolalikar (2013) find that the Pre-Conception and Pre-Natal Diagnostic Techniques Act of 1994, which was implemented in 1996 to prohibit prenatal sex detection and sex-selective abortion, had a significantly negative effect on the ratio of male to female births. Rastogi and Sharma (2022) further show that this policy has an unintended effect of reducing the educational attainment of girls. This reflects the increase in the number of unwanted daughters.[33] In addition, there is a small increase in the fertility rate by 0.02. This decreases education expenditure for sons slightly and leads the share of men with high education to drop by 1.0 pp.

5 Causes and Consequences of Missing Women

I next use the model economy as a quantitative laboratory to study the causes and consequences of missing women. I focus on economic factors, i.e., gender discrimination, dowries, and old-age support, and try to asses their importance relative to intrinsic sex preferences.

5.1 Economic Factors

To assess the effects of the economic factors on the sex ratio and their effects on fertility, education, and female labor supply, the following counterfactual experiments are conducted:

(i) I assume that daughters provide the same old-age support as sons. To this end, monetary transfers from each son and each daughter to parents given the number of sons and daughters are determined by the same function,

$$t_f(q_m, q_f) = t_m(q_m, q_f) = t_{m1}(q_m + q_f)^{-\kappa_m}.$$

[33]In line with the findings of Nandi and Deolalikar (2013) and Rastogi and Sharma (2022), Anukriti et al. (2022) find that spread of ultrasound machines in India after the middle 1980s increased the ratio of male to female births but increased parental investment in girls' health.

(ii) I eliminate dowry payments and set δ to 0.[34]

(iii) I assume that there is no labor market discrimination against women, i.e., λ to 1.

(iv) Finally, I combine (i), (ii), and (iii) and remove all the gender differences in the economic factors.

The results are reported in Table 11.

Table 11: Effects of the Gender Difference in Economic Factors

	Benchmark	Equal support	No dowry	No discrimination	No difference in economic factors
Sex ratio at birth	1.139	1.096	1.052	1.232	1.052
Fertility	3.07	2.97	3.79	2.20	2.40
Education expenditure per son	0.0570	0.0566	0.0406	0.0809	0.0688
Education expenditure per daughter	0.0379	0.0449	0.0242	0.0587	0.0558
Men with high education (%)	64.6	64.1	47.6	86.6	79.2
Women with high education (%)	49.2	55.6	31.6	75.8	72.1
Gender gap in education	15.4	8.5	15.9	10.8	7.2
Female labor (young)	0.250	0.271	0.201	0.450	0.448
Female labor (middle-aged)	0.270	0.284	0.264	0.446	0.451

The results reveal that the economic factors may increase or decrease the sex ratio and female education, depending on the trade-off between sons and daughters and the trade-off between the quantity and the quality of children. First, if the old-age support requirement is the same for sons and daughters, daughters are more valuable for parents. For this reason, fewer female fetuses are aborted, and the sex ratio at birth (SRB) decreases from 1.139 to 1.096. Moreover, education investment per daughter as a share of household income goes up from 3.8% to 4.5%, which increases the share of women with high education from 49.2% to 55.6%. Education investment per son decreases slightly, but the change is negligible.

Second, if dowry payment is prohibited, daughters are less costly. As a result, son-biased sex selections are almost not practiced, and the SRB declines to 1.052, which is close to the natural ratio of 1.05. However, parents reduce education investment in children, which has a larger effect on daughters' human capital formation. The reason is that the expected cost of each birth is reduced and parents shift their preference away

[34]This mimics the Dowry Prohibition Act 1961, although this act is difficult to implement and has not been strictly enforced.

from the quality of children to the quantity of children. As shown in the table, the fertility rate is increased to 3.79, while education investment as a share of household income is reduced by 1.6 pp for sons and 1.4 pp for daughters. This leads to a decline in the share of men and women with high education by 17.0 pp and 17.6 pp, respectively.

Third, if labor market discrimination is eliminated, the sex ratio at birth increases to 1.232, which may look surprising. One might expect that daughters become more valuable as they have higher household income, but in fact daughters' household income does not matter much for parents unless it is used to provide old-age support.[35] On the contrary, sons become more valuable, due to an increase in their wives' income. Therefore, the gender difference in providing support to parents is enhanced. Moreover, since higher labor market returns for women increase their labor supply and reduce fertility, it is more likely that parents end up with no son if they do not select their children's sex.[36] For these two reasons, parents have more incentives to select a male birth. On the other hand, higher labor market returns for women lead parents to increase education expenditure for children. More specifically, education investment increases by 2.4 pp for sons and 2.1 pp for daughters, which leads the share of men and women with high education to increase by 22.0 pp and 26.6 pp, respectively.

Finally, if all the differences in the economic factors are removed, the sex ratio at birth drops to 1.052. Education investment increases substantially, in particular, for girls. As a result, the share of men with high education increases by 14.6 pp, while the share of women with high education increases to a larger extent by 22.9 pp, which shrinks the gender education gap significantly. With an interaction of these factors, it is noteworthy that eliminating labor market discrimination does not increase the sex ratio at birth despite that women supply more labor to the market and have fewer children. The reason is that daughters become much more valuable as they provide the same support to parents as sons.

Overall, a combination of these factors can lead to an almost normal sex ratio and a

[35] For this reason, Das Gupta et al. (2003) argue that increasing female labor supply cannot reduce the sex ratio if parents still rely on sons rather than daughters for old-age support.

[36] The negative relationship between fertility and sex selection is consistent with the findings of Jayachandran (2017) and Ebenstein (2010). Jayachandran (2017) shows that the desired ratio of sons to daughters increases sharply as fertility falls in India, and Ebenstein (2010) finds that in Chinese provinces where the fines for violating the One-Child Policy were higher, the fertility rate was lower and the sex ratio was more biased.

small gender education gap, together with a large improvement in female labor supply and a substantial drop in fertility. Figures A9 and A10 in Appendix D show the PPR and the probability that the next birth is a son conditional on the sex composition of existing children. The son-biased fertility stopping rule and son-biased sex selections are almost no longer practiced.

5.2 Economic Factors vs. Intrinsic Sex Preference

Next, I compare the relative importance of the economic factors and intrinsic son preference. To this end, I start by eliminating intrinsic son preference, i.e., set ω to 1, after making the economic factors symmetric for sons and daughters. The results are reported in Table 12.[37]

Table 12: Effects of the Gender Difference in Economic Factors and Intrinsic Son Preference

	Benchmark	No difference in economic factors	No difference in economic factors + no intrinsic son preference	No intrinsic son preference
Sex ratio at birth	1.139	1.052	1.050	1.091
Fertility	3.07	2.40	2.34	3.04
Education expenditure per son	0.0570	0.0688	0.0646	0.0533
Education expenditure per daughter	0.0379	0.0558	0.0638	0.0440
Men with high education (%)	64.6	79.2	77.9	61.7
Women with high education (%)	49.2	72.1	77.4	54.1
Gender gap in education	15.4	7.2	0.5	7.6
Female labor (young)	0.250	0.448	0.459	0.265
Female labor (middle-aged)	0.270	0.451	0.461	0.290

Once the gender differences in the economic factors are removed, further eliminating intrinsic sex preference has limited effects on the SRB and the gender gap in education expenditure. First, the sex ratio is reduced by only 0.002. Second, the change in education expenditure as a share of household income is 0.8 pp for girls and -0.4 pp for boys, which

[37]Figures A11 and A12 in Appendix D show the PPR and the probability that the next birth is a son conditional on the sex composition of existing children. The son-biased fertility stopping rule and son-biased sex selections are no longer practiced.

is much smaller than the effects of the economic factors (1.8 pp and 1.2 pp). In addition, the effects on fertility and female labor supply are small.

What happens if only intrinsic son preference is removed? The last column of Table 12 shows that there is still a decrease in the SRB and a reduction in the gender education gap. However, its effects on the SRB and education expenditure are smaller than the effects of removing only the gender differences in the economic factors. In addition, its effects on fertility and female labor supply are minor. The results suggest that the preference for sons can be largely attributed to the gender differences in the economic factors.

6 Policies to Normalize the Sex Ratio

How can governments compact imbalanced sex ratios? To answer this question, I first compare a subsidy for female births and a subsidy for female education, two measures that are widely taken by Indian governments. Next, I examine how a pension system can affect relevant household behaviors.

6.1 Subsidize Female Births vs. Subsidize Female Education

Over years, Indian governments have introduced various conditional cash transfer programs to combat the biased sex ratio. Most policies aimed at normalizing the sex ratio at birth and improving the educational attainment of girls, while some also attempted to increase health investment for girls and delay their marriage.[38] Two important measures are cash transfers to parents with female births and cash transfers to parents whose daughters complete some levels of education, such as primary or secondary education. Which policy is more effective in normalizing the sex ratio? Which policy can increase education investment, especially for girls? How do they affect fertility and female labor supply?

In this subsection, I compare the impacts of these two policies. To this end, I introduce either a lump-sum subsidy for female births or a proportional subsidy for female education. For the former, I assume that parents can receive a transfer amounting to 0.02 for each female birth, which is equivalent to 40% of the annual income of a male without secondary

[38] See Anukriti (2018) and Biswas et al. (2021) for an overview of such policies.

education. For the latter, I assume that 20% of education expenditure for daughters is paid by the government.[39] Both subsidies are financed by a proportional labor income tax. The outcome variables are reported in Table 13.

Table 13: Effects of the Subsidies for Female Births and Female Education

	Benchmark	Birth subsidy	Education subsidy
Sex ratio at birth	1.139	1.063	1.091
Fertility	3.07	3.55	2.82
Education expenditure per son	0.0570	0.0478	0.0556
Education expenditure per daughter	0.0379	0.0303	0.0568
Men with high education (%)	64.6	54.0	65.0
Women with high education (%)	49.2	38.5	72.7
Gender gap in education	15.4	15.5	-7.7
Female labor (young)	0.250	0.207	0.305
Female labor (middle-aged)	0.270	0.284	0.315
Labor income tax rate	0	0.014	0.010

Both policies are effective in reducing the ratio of male to female births. The SRB is reduced to 1.063 if female births are subsidized, and 1.091 if female education is subsidized. The underlying mechanisms, however, are different. Subsidizing female births directly reduces the fixed cost of daughters, while subsidizing female education makes education expenditure for daughters cheaper from the perspective of parents, therefore, increasing the value of having daughters.

Despite similar effects on sex selection, the two policies have different effects on female education, because parental quantity-quality trade-off is shifted into opposite directions. By reducing the expected cost of birth, the subsidy for female births increases fertility (to 3.55) and decreases education investment. As a result, the share of men and women with high education is reduced to 54.0% and 38.5%, respectively.[40] In contrast, the subsidy for female education directly motivates parents to increase daughters' education expenditure, which increases the share of women with high education to 72.7%. Moreover, this leads

[39]These subsidy rates are roughly comparable to some policies in India, such as the transfer for female births of the *Ladli* Scheme in Delhi from 2008 and the transfer for female education of *Apni Beti Apna Dhan* in Haryana in 1994-1998 and *Dhanlakshmi*. In Appendix D, I also evaluate the impact of alternative subsidy amounts and show that the basic findings are largely unchanged.

[40]Now household income is net of taxes and includes transfers, so it makes little sense to compare education expenditure as a share of household income. Therefore, I look at the share of individuals with high education.

to a lower fertility rate (2.82), which in turn increases education expenditure per son. For this reason, although parents shift some resources away from sons to daughters, the share of men with high education is almost not changed (65.0%).

Finally, the two policies can affect female labor supply differently through two channels, namely fertility and education. First, a higher fertility rate implies less time for market work. Second, more educated women supply more labor in the market due to a higher wage rate. As a result, subsidizing female births decreases young women's labor supply by 4.3 pp of their time, while subsidizing female education can increase it by 5.5 pp.

To sum up, due to the quantity-quality trade-off, although a female birth subsidy can normalize the SRB, it has unfavorable effects on fertility, education, and female labor supply. In contrast, a female education subsidy can lead to improvements in the SRB, fertility, female education, and female labor supply.

6.2 Pension System

As shown in the previous section, concern for old-age support is an important reason for sex selection. The role of children for old-age support, however, will be weakened if parents are covered by a pension system. In particular, the gender difference for old-age support will be less important. How much could the pension system affect parental sex preference? How could it affect other outcomes, including fertility, education investment, and female labor supply?

To answer these questions, a pay-as-you-go pension system is introduced into the model. I assume that men receive pension payment, which is financed by a labor income tax on currently working men. Since women supply much less labor than men, it is assumed that they will not receive pensions and that they do not pay taxes. The replacement rate is set to 50%, which is the targeted replacement rate of the Employees Provident Fund Scheme, the largest pension system in India (Kim and Bhardwaj, 2011).[41,42] The results are shown in Table 14.

[41] According to the 2011 census, only about 12% of the workforce was covered under various pension schemes. See Kim and Bhardwaj (2011) and OECD (2020) for an overview of the pension schemes in India.

[42] See more results in Table A9 in Appendix D with different replacement rates.

Table 14: Effects of the Pension System

	Benchmark	Only pension	Half support from children	No Support from children
Sex ratio at birth	1.139	1.092	1.084	1.077
Fertility	3.07	3.03	3.04	3.02
Education expenditure per son	0.0570	0.0540	0.0525	0.0511
Education expenditure per daughter	0.0379	0.0384	0.0391	0.0405
Men with high education (%)	64.6	56.6	55.3	54.0
Women with high education (%)	49.2	44.4	44.9	46.1
Gender gap in education	15.4	12.2	10.4	8.0
Female labor (young)	0.250	0.295	0.296	0.299
Female labor (middle-aged)	0.270	0.279	0.288	0.298
Savings rate (middle-aged)	0.361	0.287	0.294	0.301
Labor income tax rate	0	0.141	0.140	0.140

To finance pension, the tax rate is about 14%. Since children become less valuable, parents now have slightly fewer children (-0.04). Although education expenditure as a fraction of household income does not change much, its absolute value is reduced substantially, since household income net of tax is reduced. As a result, the share of men with high education is reduced to 56.6% and it is reduced to 44.4% for women. Since the gender difference in old-age support is less important, there is also a reduction in the SRB by 0.047 to 1.092.[43]

So far the analysis has assumed that the pension system does not change the old-age support that the parents received from their children. Yet after the pension system is introduced, there may also be a reduction in the old-age support requirement, which may happen in the long run as a cultural change (Cheng et al., 2018; Mukherjee, 2020). To examine this possibility, I consider two hypothetical cases, where the old-aged support is cut by half or full. Now the SRB further drops to 1.084 and 1.077, respectively. The educational attainment of men is further reduced, but women receive more education. Suppose that the old-age support requirement is cut by half, then the share of women

[43] The model predictions are consistent with the empirical findings in the literature. First, Boldrin et al. (2015) find that the fertility rate is lower in countries with higher social security taxes. Bau (2021) finds that pension plans reduced parental investment in children's education in Indonesia and Ghana, where parents rely on either sons or daughters for old-age support. Third, Ebenstein and Leung (2010) exploit the pension plan introduced in rural China in the 1990s and show that access to this program reduced son-biased sex selections.

with high education increases by 0.5 pp, compared with the case with pension but without any change in old-age support requirement. The reason for this improvement is that as the old-age support motive for educating sons is further reduced, parents shift more resources to daughters.

To sum up, the pension system can reduce the SRB, but it hurts education investment for both sons and daughters. If the pension system is accompanied by a reduction in old-age support requirements, the negative effect on girls' education can be alleviated.

7 Conclusions

In developing countries, a preference for sons is common, which is manifested as parents desiring more sons than daughters and caring more about sons' education than daughters'. However, the relative importance of different factors for son preference is not well understood, nor are its effects on fertility and female labor supply. In this study, an overlapping-generation model of fertility, education investment, and marriage formation is built to quantify the causes and consequences of missing women in India.

The quantitative analysis reveals that economic factors play a more important role than intrinsic sex preference in generating a preference for sons. In the benchmark model, the SRB is 1.139. 64.6% of men complete secondary education, while only 49.2% of women achieve this. If there is no gender difference in the economic factors, namely old-age support, marriage cost, and labor market discrimination, the SRB drops by 0.087 to 1.052, which is very close to the natural ratio. The share of men with secondary education rise to 79.2%, and the share of women with secondary education rises by a larger degree to 72.1%. Further eliminating intrinsic son preference has a very limited effect on the SRB by reducing it to 1.050. It decreases the share of men with secondary education slightly to 77.9% but increases the share of women with secondary education to 77.4%.

Economic factors also have larger effects on fertility and female labor supply than intrinsic son preference. If the economic factors become symmetric for sons and daughters, the fertility rate drops from 3.07 in the benchmark economy to 2.40. The labor supply of young women increases from 25.0% of their time endowment in the benchmark economy to 44.8%, and for middle-aged women, it rises from 27.0% to 45.1%. If intrinsic son preference is further removed, there will be another small drop in fertility to 2.34. Young women supply 45.9% of their time, and middle-aged women would supply 46.1%.

Finally, the model is used to evaluate some policies that may affect parental sex preference. A comparison is made between subsidizing female births and subsidizing female education. Although both policies are effective in reducing the SRB, they have different effects on fertility, education, and female labor supply by shifting the parental quantity-quality trade-off into opposite directions. More specifically, a subsidy for female births amounting to 40% of the annual income of a male with less than secondary education reduces the SRB to 1.063. But it raises the fertility rate to 3.55, reduces the share of men and women with secondary education to 54.0% and 38.5%, respectively, and reduces young women's labor supply to 20.7% of their time endowment. In contrast, a subsidy of 20% of education expenditure for girls reduces the SRB to 1.091. Moreover, it reduces the fertility rate to 2.82, increases the share of men with secondary education slightly to 65.0% and the share of women with secondary education substantially to 72.7%, and increases young women's labor supply to 31.0% of their time.

Finally, introducing a pension system reduces sex selection, fertility, and children's education. A pay-as-you-go pension system with a 50% replacement rate reduces the SRB to 1.092, reduces the fertility rate by 0.04 to 3.03, and reduces the share of men and women with secondary education to 56.6% and 44.4%, respectively. If the pension system leads to a reduction in the old-age support requirement of children, the effect on the sex ratio will be slightly larger. It further decreases men's educational attainment but increases women's educational attainment.

References

Abrevaya, J. (2009). Are there missing girls in the United States? Evidence from birth data. *American Economic Journal: Applied Economics*, *1*(2), 1–34.

Adams, J. J. (2021). Urbanization, long-run growth, and the demographic transition. *Journal of Demographic Economics*, 1–47.

Agrawal, T., & Agrawal, A. (2019). Who gains more from education? A comparative analysis of business, farm and wage workers in India. *Journal of Development Studies*, *55*(6), 1081–1098.

Almond, D., Edlund, L., & Milligan, K. (2013). Son preference and the persistence of culture: Evidence from South and East Asian immigrants to Canada. *Population and Development Review*, *39*(1), 75–95.

Altindag, O. (2016). Son preference, fertility decline, and the nonmissing girls of Turkey. *Demography*, *53*(2), 541–566.

Anderson, S. (2007). The economics of dowry and brideprice. *Journal of Economic Perspectives*, *21*(4), 151–174.

Andrews, I., Gentzkow, M., & Shapiro, J. M. (2017). Measuring the sensitivity of parameter estimates to estimation moments. *Quarterly Journal of Economics*, *132*(4), 1553–1592.

Anukriti, S. (2018). Financial incentives and the fertility-sex ratio trade-off. *American Economic Journal: Applied Economics*, *10*(2), 27–57.

Anukriti, S., Bhalotra, S., & Tam, E. H. (2022). On the quantity and quality of girls: Fertility, parental investments and mortality. *Economic Journal*, *132*(641), 1–36.

Arnold, F. (1985). Measuring the effect of sex preference on fertility: The case of Korea. *Demography*, 280–288.

Arnold, F., Kishor, S., & Roy, T. K. (2002). Sex-selective abortions in India. *Population and Development Review*, *28*(4), 759–785.

Bairagi, R., & Langsten, R. L. (1986). Sex preference for children and its implications for fertility in rural Bangladesh. *Studies in Family Planning*, *17*(6), 302–307.

Bar, M., & Leukhina, O. (2010). Demographic transition and industrial revolution: A macroeconomic investigation. *Review of Economic Dynamics*, *13*(2), 424–451.

Barcellos, S. H., Carvalho, L. S., & Lleras-Muney, A. (2014). Child gender and parental investments in India: Are boys and girls treated differently? *American Economic Journal: Applied Economics*, *6*(1), 157–89.

Bardhan, P. K. (1974). On life and death questions. *Economic and Political Weekly*, 1293–1304.

Barro, R. J., & Becker, G. S. (1989). Fertility choice in a model of economic growth. *Econometrica*, 481–501.

Bau, N. (2021). Can policy change culture? Government pension plans and traditional kinship practices. *American Economic Review*, *111*(6), 1880–1917.

Becker, G. S. (1981). A treatise on the family. *Cambridge: Harvard University*.

Behrman, J. R., Foster, A. D., Rosenweig, M. R., & Vashishtha, P. (1999). Women's schooling, home teaching, and economic growth. *Journal of Political Economy*, *107*(4), 682–714.

Beltrán Tapia, F. J., & Raftakis, M. (2021). Sex ratios and gender discrimination in Modern Greece. *Population Studies*, 1–18.

Bhalotra, S., Chakravarty, A., & Gulesci, S. (2020). The price of gold: Dowry and death in India. *Journal of Development Economics*, *143*, 102413.

Bhalotra, S., Clots-Figueras, I., & Iyer, L. (2021). Religion and abortion: The role of politician identity. *Journal of Development Economics*, *153*, 102746.

Bhalotra, S. R., & Cochrane, T. (2010). Where have all the young girls gone? Identification of sex selection in India. *Working paper*.

Biswas, N., Cornwell, C., & Zimmermann, L. V. (2021). The power of Lakshmi: Monetary incentives for raising a girl. *GLO discussion paper*.

Blau, F. D., Kahn, L. M., Brummund, P., Cook, J., & Larson-Koester, M. (2020). Is there still son preference in the United States? *Journal of Population Economics*, *33*(3), 709–750.

Boldrin, M., De Nardi, M., & Jones, L. E. (2015). Fertility and social security. *Journal of Demographic Economics*, *81*(3), 261–299.

Bongaarts, J. (2013). The implementation of preferences for male offspring. *Population and Development Review*, *39*(2), 185–208.

Boserup, E. (1970). *Woman's role in economic development*. New York: St. Martin's.

Bourne, K. L., & Walker, G. M., Jr. (1991). The differential effect of mothers' education on mortality of boys and girls in India. *Population Studies*, *45*(2), 203–219.

Calvi, R., & Keskar, A. (2021). Dowries, resource allocation, and poverty. *Journal of Economic Behavior & Organization*, *192*, 268–303.

Carlos Córdoba, J., & Ripoll, M. (2019). The elasticity of intergenerational substitution, parental altruism, and fertility choice. *Review of Economic Studies*, *86*(5), 1935–1972.

Carol, S., & Hank, K. (2020). Natives' and immigrants' gender preferences for children in Germany. *European Journal of Population*, *36*(2), 235–246.

Carranza, E. (2014). Soil endowments, female labor force participation, and the demographic deficit of women in India. *American Economic Journal: Applied Economics*, *6*(4), 197–225.

Caucutt, E. M., Guner, N., & Rauh, C. (2021). Is marriage for white people? Incarceration, unemployment, and the racial marriage divide. *Working paper*.

Cavalcanti, T., Kocharkov, G., & Santos, C. (2021). Family planning and development: Aggregate effects of contraceptive use. *Economic Journal*, *131*(634), 624–657.

Cavalcanti, T., & Tavares, J. (2016). The output cost of gender discrimination: A model-based macroeconomics estimate. *Economic Journal*, *126*(590), 109–134.

Chao, F., Gerland, P., Cook, A. R., & Alkema, L. (2019). Systematic assessment of the sex ratio at birth for all countries and estimation of national imbalances and regional reference levels. *Proceedings of the National Academy of Sciences*, *116*(19), 9303–9311.

Cheng, L., Liu, H., Zhang, Y., & Zhao, Z. (2018). The heterogeneous impact of pension income on elderly living arrangements: Evidence from China's new rural pension scheme. *Journal of Population Economics*, *31*(1), 155–192.

Chiplunkar, G., & Weaver, J. (2021). Marriage markets and the rise of dowry in India. *Working paper*.

Choi, E. J., & Hwang, J. (2015). Child gender and parental inputs: No more son preference in Korea? *American Economic Review*, *105*(5), 638–43.

Chung, W., & Das Gupta, M. (2007). The decline of son preference in South Korea: The roles of development and public policy. *Population and Development Review*, *33*(4), 757–783.

Clark, S. (2000). Son preference and sex composition of children: Evidence from India. *Demography*, *37*(1), 95–108.

Dahl, G. B., & Moretti, E. (2008). The demand for sons. *Review of Economic Studies*, *75*(4), 1085–1120.

Das Gupta, M. (2019). Is banning sex-selection the best approach for reducing prenatal discrimination? *Asian Population Studies*, *15*(3), 319–336.

Das Gupta, M., Jiang, Z., Li, B., Xie, Z., Chung, W., & Bae, H.-O. (2003). Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea. *Journal of Development Studies*, *40*(2), 153–187.

Davies, J. B., & Zhang, J. (1997). The effects of gender control on fertility and children's consumption. *Journal of Population Economics*, *10*(1), 67–85.

de la Croix, D., & Doepke, M. (2003). Inequality and growth: Why differential fertility matters. *American Economic Review*, *93*(4), 1091–1113.

Delventhal, M. J., Fernández-Villaverde, J., & Guner, N. (2021). Demographic transitions across time and space. *Working paper*.

Deshpande, A., Goel, D., & Khanna, S. (2018). Bad karma or discrimination? Male–female wage gaps among salaried workers in India. *World Development*, *102*, 331–344.

Dharmalingam, A., Rajan, S., & Morgan, S. P. (2014). The determinants of low fertility in India. *Demography*, *51*(4), 1451–1475.

Diamond-Smith, N., Luke, N., & McGarvey, S. (2008). 'Too many girls, too much dowry': Son preference and daughter aversion in rural Tamil Nadu, India. *Culture, Health and Sexuality*, *10*(7), 697–708.

Dubuc, S., & Coleman, D. (2007). An increase in the sex ratio of births to India-born mothers in England and Wales: Evidence for sex-selective abortion. *Population and Development Review*, *33*(2), 383–400.

Dyson, T. (2012). Causes and consequences of skewed sex ratios. *Annual Review of Sociology*, *38*, 443–461.

Ebenstein, A. (2010). The "missing girls" of China and the unintended consequences of the one child policy. *Journal of Human Resources*, *45*(1), 87–115.

Ebenstein, A. (2021). Elderly coresidence and son preference: Can pension reforms solve the 'missing women' problem? *Working paper*.

Ebenstein, A., & Leung, S. (2010). Son preference and access to social insurance: Evidence from China's rural pension program. *Population and Development Review*, *36*(1), 47–70.

Eckstein, Z., Keane, M., & Lifshitz, O. (2019). Career and family decisions: Cohorts born 1935–1975. *Econometrica*, *87*(1), 217–253.

Edlund, L. (2000). The marriage squeeze interpretation of dowry inflation: A comment. *Journal of Political Economy*, *108*(6), 1327–1333.

Fernandez, R., Guner, N., & Knowles, J. (2005). Love and money: A theoretical and empirical analysis of household sorting and inequality. *Quarterly Journal of Economics*, *120*(1), 273–344.

Fernández, R., & Rogerson, R. (2001). Sorting and long-run inequality. *Quarterly Journal of Economics*, *116*(4), 1305–1341.

Fuse, K. (2010). Variations in attitudinal gender preferences for children across 50 less-developed countries. *Demographic Research*, *23*, 1031–1048.

Galor, O., & Weil, D. N. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American Economic Review*, *90*(4), 806–828.

Ganatra, B., & Hirve, S. (2002). Induced abortions among adolescent women in rural Maharashtra, India. *Reproductive Health Matters*, *10*(19), 76–85.

Goel, P. A., & Barua, R. (2021). Female education, marital assortative mating, and dowry: Theory and evidence from districts of India. *Journal of Demographic Economics*, 1–27.

Goodkind, D. (1996). On substituting sex preference strategies in East Asia: Does prenatal sex selection reduce postnatal discrimination? *Population and Development Review*, 111–125.

Greenwood, J., & Seshadri, A. (2002). The US demographic transition. *American Economic Review*, *92*(2), 153–159.

Guilmoto, C. Z. (2012). Skewed sex ratios at birth and future marriage squeeze in China and India, 2005–2100. *Demography*, *49*(1), 77–100.

Hansen, G. D., & Prescott, E. C. (2002). Malthus to solow. *American Economic Review*, *92*(4), 1205–1217.

Hesketh, T., & Zhu, W. (2006). Abnormal sex ratios in human populations: Causes and consequences. *Proceedings of the National Academy of Sciences*, *103*(36), 13271–13275.

Jayachandran, S. (2017). Fertility decline and missing women. *American Economic Journal: Applied Economics*, *9*(1), 118–39.

Jiang, Q., Feldman, M. W., & Li, S. (2014). Marriage squeeze, never-married proportion, and mean age at first marriage in China. *Population Research and Policy Review*, *33*(2), 189–204.

Kalemli-Ozcan, S. (2002). Does the mortality decline promote economic growth? *Journal of Economic Growth*, *7*(4), 411–439.

Kalemli-Ozcan, S. (2003). A stochastic model of mortality, fertility, and human capital investment. *Journal of Development Economics*, *70*(1), 103–118.

Kim, C., & Bhardwaj, G. (2011). *South Asia pension forum: Fostering inclusive and sustainable pension systems in the region* (Vol. 26). Asian Development Bank.

Kumar, S., & Sinha, N. (2020). Preventing more "missing girls": A review of policies to tackle son preference. *The World Bank Research Observer*, *35*(1), 87–121.

Leung, S. F. (1994). Will sex selection reduce fertility? *Journal of Population Economics*, *7*(4), 379–392.

Maertens, A., & Chari, A. (2020). What's your child worth? An analysis of expected dowry payments in rural India. *World Development*, *130*, 104927.

Marco-Gracia, F. J., & Beltrán Tapia, F. J. (2021). Son preference, gender discrimination, and missing girls in rural Spain, 1750–1950. *Population and Development Review*, *47*(3), 665–689.

Mukherjee, A. (2020). Intergenerational altruism and retirement transfers: Evidence from the social security notch. *Journal of Human Resources*, 0419–10140R3.

Nandi, A., & Deolalikar, A. B. (2013). Does a legal ban on sex-selective abortions improve child sex ratios? Evidence from a policy change in India. *Journal of Development Economics*, *103*, 216–228.

OECD. (2020). *Pensions at a glance 2019: OECD and G20 indicators*. Paris: OECD Publishing.

Orzack, S. H., Stubblefield, J. W., Akmaev, V. R., Colls, P., Munné, S., Scholl, T., Steinsaltz, D., & Zuckerman, J. E. (2015). The human sex ratio from conception to birth. *Proceedings of the National Academy of Sciences*, *112*(16), E2102–E2111.

Pande, R. P., & Astone, N. M. (2007). Explaining son preference in rural India: The independent role of structural versus individual factors. *Population Research and Policy Review*, *26*(1), 1–29.

Powell-Jackson, T., Mazumdar, S., & Mills, A. (2015). Financial incentives in health: New evidence from India's Janani Suraksha Yojana. *Journal of Health Economics*, *43*, 154–169.

Qian, N. (2008). Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. *Quarterly Journal of Economics*, *123*(3), 1251–1285.

Rao, V. (1993). The rising price of husbands: A hedonic analysis of dowry increases in rural India. *Journal of Political Economy*, *101*(4), 666–677.

Rastogi, G., & Sharma, A. (2022). Unwanted daughters: The unintended consequences of a ban on sex-selective abortions on the educational attainment of women. *Journal of Population Economics*, 1–44.

Rosenblum, D. (2013). The effect of fertility decisions on excess female mortality in India. *Journal of Population Economics*, *26*(1), 147–180.

Rosenzweig, M. R., & Schultz, T. P. (1982). Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India. *American Economic Review*, *72*(4), 803–815.

Sandström, G., & Vikström, L. (2015). Sex preference for children in German villages during the fertility transition. *Population Studies*, *69*(1), 57–71.

Sen, A. (1990). More than 100 million women are missing. *New York Review of Books*, *37*(20), 61–66.

Seshadri, A., & Zhou, A. (2022). Intergenerational mobility begins before birth. *Journal of Monetary Economics*.

Sinha, N., & Yoong, J. (2009). Long-term financial incentives and investment in daughters: Evidence from conditional cash transfers in North India.

Trivers, R. L., & Willard, D. E. (1973). Natural selection of parental ability to vary the sex ratio of offspring. *Science*, *179*(4068), 90–92.

Williamson, N. E. (1976). *Sons or daughters: A cross-cultural survey of parental preferences*. Sage Publications.

Appendices

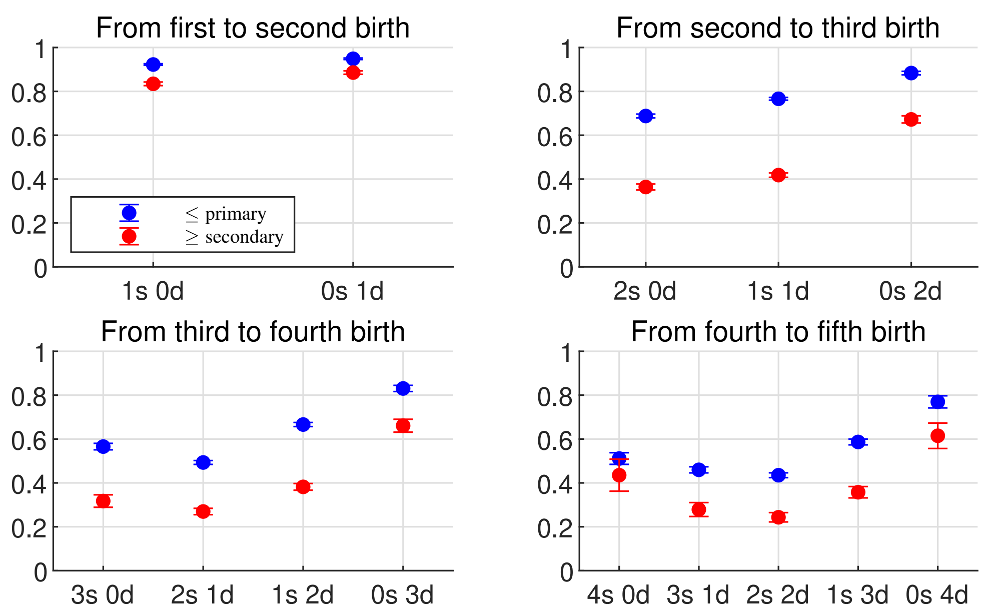
A Additional Empirical Facts

Extensive robustness checks are conducted for the facts documented in the main text. The analysis is repeated for different groups of people in the same data, using different waves of the same survey, and for other countries.

Fact 1. Son-biased Fertility Stopping Rule.

In the main text, this fact is documented using the sample of women aged 40-49 with no more than 6 children in the Indian DHS in 2015-2016.

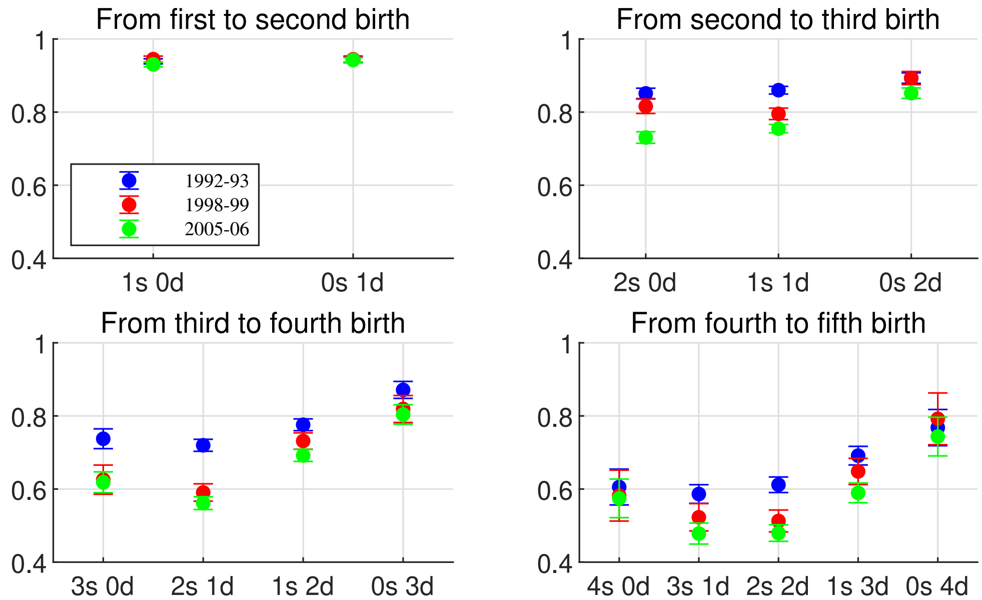
First, the analysis is repeated separately for women with and without secondary education. The results are plotted in Figure A1, suggesting that a son-biased fertility stopping rule is practiced by both groups.



Notes. Data are from the Indian Demographic and Health Survey in 2015-2016.

Figure A1. Parity Progression Ratio Conditional on the Sex Composition of Previous Children by Women's Education.

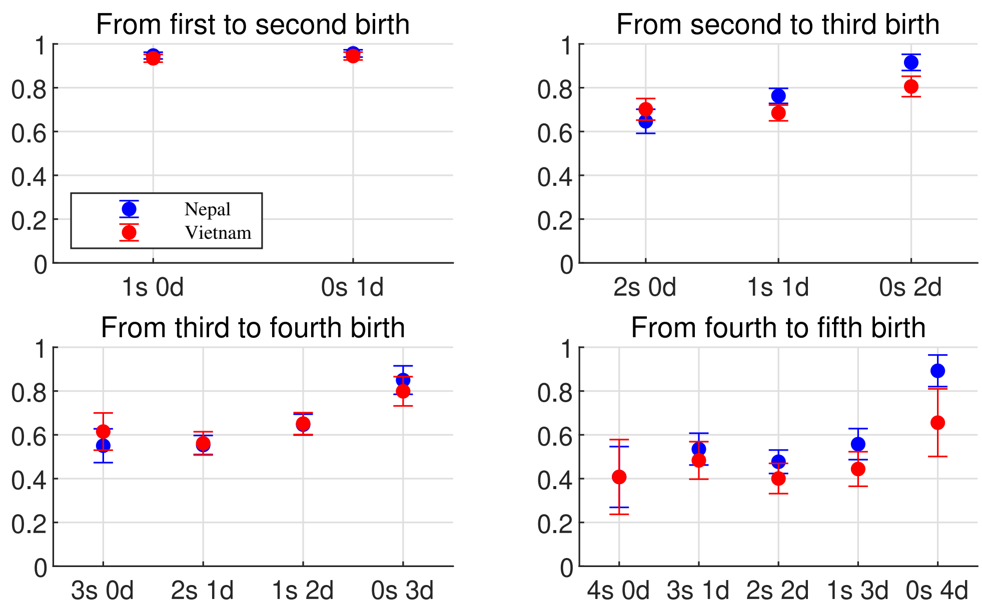
Second, the analysis is repeated for the previous waves of the survey in 1992-1993, 1998-1999, and 2005-2006. The results are plotted in Figure A2, suggesting that a son-biased fertility stopping rule has been persistent in India.



Notes. Data are from the Indian Demographic and Health Survey in 1992-1993, 1998-1999, and 2005-2006.

Figure A2. Parity Progression Ratio Conditional on the Sex Composition of Previous Children in Previous Surveys.

Finally, the analysis is repeated for Nepal and Vietnam, where survey evidence indicates that women wanted more sons than daughters (Fuse, 2010). I use data from the Nepalese DHS in 2016 and the Vietnamese DHS in 2002. The results are plotted in Figure A3. Women in these two countries also practice a son-biased fertility stopping rule.



Notes. Data are from the Nepalese DHS in 2016 and the Vietnamese DHS in 2002.

Figure A3. Parity Progression Ratio Conditional on the Sex Composition of Previous Children in Nepal and Vietnam.

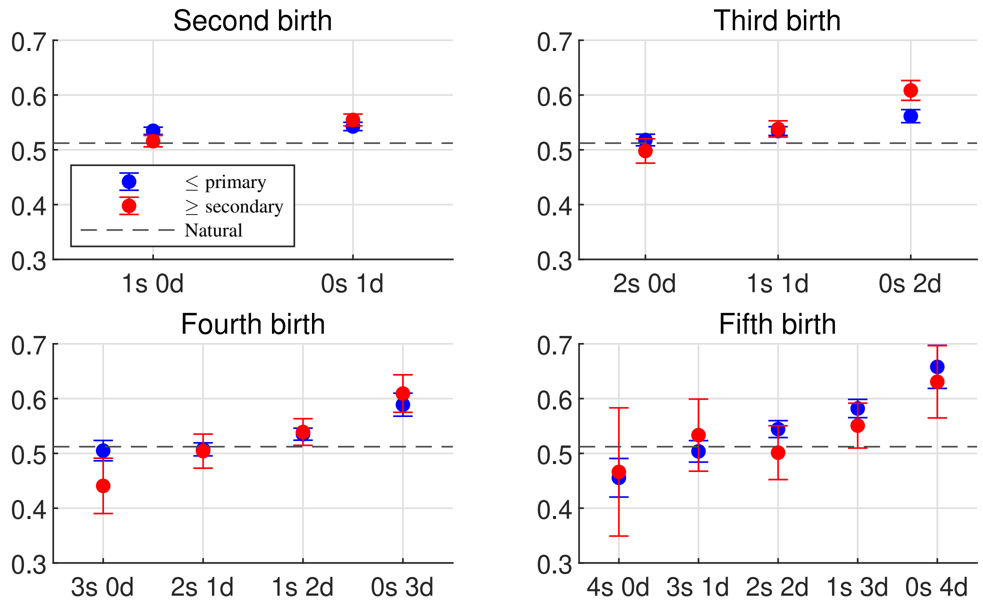
Fact 2. Son-biased Sex Selection.

In the main text, this fact is documented using the sample of women aged 40-49 with no more than 6 children in the Indian DHS in 2015-2016.

First, the analysis is repeated for women with different education. The results are plotted in Figure A4, suggesting son-biased sex selections were practiced by both groups of women.

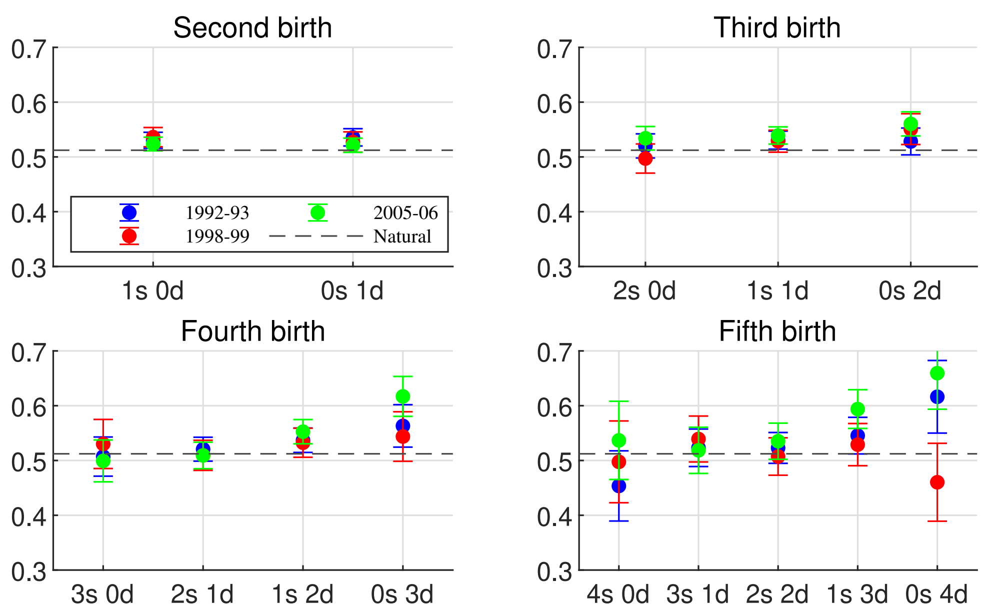
Second, the analysis is repeated for the previous waves of the survey in 1992-1993, 1998-1999, and 2005-2006. The results are shown in Figure A5, suggesting that sex selections were also practiced by women in the previous waves. One difference is that parents in these waves were less likely to select their children's sex. The reason might be that they had less access to sex selection technologies since ultrasound machines were not introduced to India until the middle 1980s and were not widely accessible until the middle 1990s (Bhalotra and Cochrane, 2010).

Finally, the analysis is repeated for Nepal and Vietnam. The results are plotted in Figure A6. There is weak evidence that in Nepal the second and third birth were more likely to be a boy if women had no sons. But no systematic pattern is observed for Vietnam. The results suggest that son-biased sex selection may or may not appear in



Notes. Data are from the Indian DHS in 2015-16.

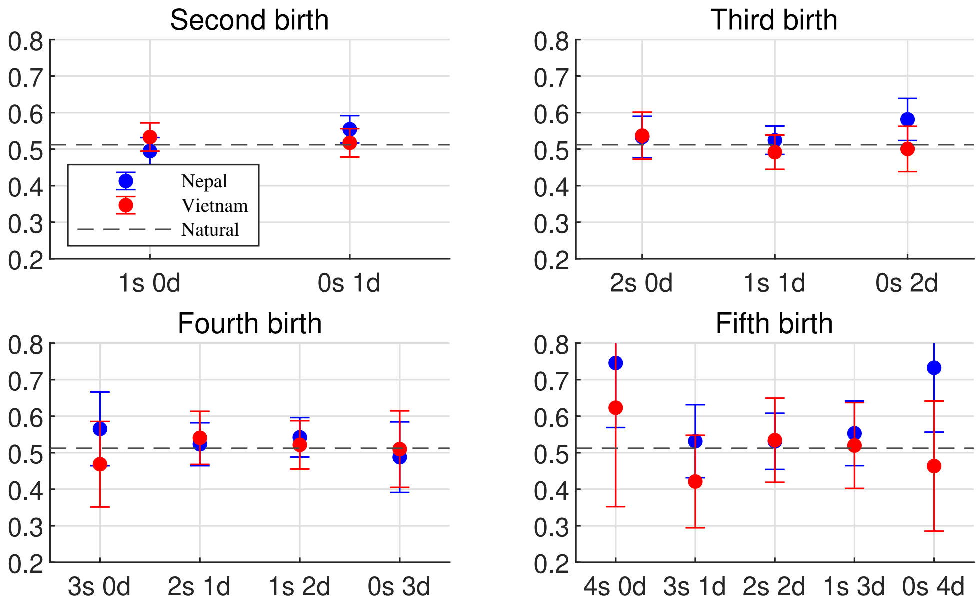
Figure A4. Probability of Having a Son Conditional on the Sex Composition of Previous Children by Women's Education



Notes. Data are from the Indian DHS in 1992-1993, 1998-1999, and 2005-2006.

Figure A5. Probability of Having a Son Conditional on the Sex Composition of Previous Children in Previous Surveys.

countries with reported son preference, since it depends on both the desire of parents and their constraints (Bongaarts, 2013).



Notes. Data are from the Nepalese DHS in 2016 and the Vietnamese DHS in 2002.

Figure A6. Probability of Having a Son Conditional on the Sex Composition of Previous Children in Nepal and Vietnam.

Fact 3 and 4. Quantity-Quality Trade-off and Gender Education Gap.

In the main text, these two facts are based on the sample of men and women aged 23-32 in the Indian HDS in 2011-2012. I focus on this age group since the socioeconomic environments, which affect household decisions, can be different for people born in different years. To check the robustness of the results, the analysis is repeated for those aged 33-42 and 43-52. The results are reported in Table A1, suggesting that the two facts can also be observed for other age groups.

Table A1: Family Size, Gender and Years of Schooling for Other Age Groups

	Aged 33-42			Aged 43-52		
	(1)	(2)	(3)	(4)	(5)	(6)
	Male	Female	Both	Male	Female	Both
Number of male children	-0.200***	-0.262***	-0.228***	-0.100**	-0.113***	-0.105***
	(0.034)	(0.042)	(0.027)	(0.041)	(0.042)	(0.030)
Number of female children	-0.065**	-0.038	-0.048*	-0.084**	0.048	-0.014
	(0.031)	(0.039)	(0.024)	(0.036)	(0.041)	(0.028)
Male			3.073***			3.060***
			(0.080)			(0.100)
Father's years of schooling	0.366***	0.425***	0.396***	0.388***	0.525***	0.465***
	(0.014)	(0.013)	(0.010)	(0.020)	(0.018)	(0.014)
Mother's years of schooling	0.544***	0.249***	0.403***	0.628***	0.303***	0.449***
	(0.019)	(0.017)	(0.013)	(0.033)	(0.028)	(0.021)
Birth year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	12534	11657	24191	6981	9857	16838
R^2	0.389	0.269	0.379	0.391	0.258	0.353

Notes. Standard errors in parentheses. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Fact 5. Women's Education and Fertility.

In the main text, this fact is based on data from the Indian DHS in 2015-2016. Table A2 shows that the negative relationship between women's education and fertility also exists in the previous waves of the survey in India in 1992-1993, 1998-1999, and 2005-2006. Table A3 shows that it can be observed in Nepal and Vietnam.

Table A2: Education and Fertility in Previous Surveys

	(1)	(2)	(3)
	India 1992-93	India 1998-99	India 2005-06
Wife's years of schooling	-0.155***	-0.134***	-0.141***
	(0.005)	(0.006)	(0.003)
Husband's years of schooling	-0.026***	-0.026***	-0.024***
	(0.004)	(0.005)	(0.003)
Wife's birth year fixed effects	Yes	Yes	Yes
Husband's birth year fixed effects	Yes	Yes	Yes
N	17757	9618	21685
R^2	0.110	0.121	0.171

Notes. Sample women are aged 40-49 years. Standard errors in parentheses. *** $p < 0.01$.

Table A3: Education and Fertility in Other Countries

	(1)	(2)
	Nepal 2016	Vietnam 2002
Wife's years of schooling	-0.135***	-0.172***
	(0.012)	(0.014)
Husband's years of schooling	-0.077***	-0.070***
	(0.010)	(0.014)
Wife's birth year fixed effects	Yes	Yes
Husband's birth year fixed effects	Yes	Yes
N	2170	1654
R^2	0.180	0.315

Notes. Sample women are aged 40-49 years. Standard errors in parentheses. *** $p < 0.01$.

B Sensitivity Analysis of Estimation

This appendix analyzes how the moments affect the parameters using the sensitivity measure proposed by Andrews et al. (2017). In Table A4, each cell shows the percentage change of the parameter when the corresponding moment changes by one percent.

Table A4: Sensitivity of Parameters to Moments

	β	ν_1	ν_2	μ_α	σ_α	ω	$1/\epsilon$	μ_ξ	σ_ξ	η	e_0	z	τ
Savings rate (middle-aged)	**0.6**	0.0	0.2	0.1	-0.1	0.1	-0.3	-0.1	**-0.9**	0.0	0.2	0.0	0.0
Women's working hours (young)	0.0	**-0.6**	0.0	0.1	0.1	0.0	**0.8**	0.1	0.4	0.0	-0.1	-0.1	0.0
Women's working hours (middle-aged)	0.0	0.0	**-0.5**	0.0	0.0	0.0	0.0	0.0	**0.1**	0.0	0.0	0.0	0.0
PPR if 1s 0d	0.0	0.0	0.0	-0.1	0.0	0.0	**0.2**	0.0	0.0	0.0	**-0.1**	0.0	0.0
PPR if 0s 1d	0.0	0.0	0.0	0.0	0.0	0.0	**0.0**	0.0	**0.0**	0.0	0.0	0.0	0.0
PPR if 2s 0d	0.1	0.0	0.0	0.0	0.0	-0.1	**0.7**	0.0	**-1.1**	0.1	0.0	-0.1	-0.1
PPR if 1s 1d	0.1	**-0.4**	0.0	-0.1	-0.1	0.0	**-1.6**	0.0	-0.2	-0.1	0.1	0.1	0.0
PPR if 0s 2d	-0.1	-0.1	-0.1	-0.1	0.0	-0.1	-0.1	0.0	**-0.2**	0.0	-0.1	**-0.1**	0.0
PPR if 3s 0d	-0.1	-0.1	0.0	-0.1	0.1	0.0	**3.8**	0.0	**0.5**	0.0	-0.1	-0.1	0.0
PPR if 2s 1d	0.0	0.1	0.0	0.0	0.0	0.0	**-1.1**	0.0	**-0.3**	0.1	0.1	0.0	-0.1
PPR if 1s 2d	-0.1	-0.1	0.0	**-0.2**	0.2	0.0	-0.1	0.0	0.1	-0.1	**-0.3**	-0.1	0.1
PPR if 0s 3d	**0.1**	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	**-0.3**	0.0	0.1	0.1	0.0
PPR if 4s 0d	0.0	0.0	0.0	0.0	0.0	0.0	**3.1**	0.0	**0.5**	0.0	0.2	0.3	0.0
PPR if 3s 1d	-0.1	0.0	0.0	-0.1	0.0	0.0	**-0.2**	0.0	0.0	0.0	-0.2	**-0.2**	0.0
PPR if 2s 2d	-0.1	0.1	0.0	-0.1	-0.1	0.0	**-0.5**	0.1	**0.7**	0.0	-0.1	-0.1	-0.1
PPR if 1s 3d	0.1	0.0	0.0	-0.1	0.0	0.0	**0.3**	0.0	0.0	0.0	0.0	**0.2**	0.0
PPR if 0s 4d	0.1	0.1	0.0	0.1	0.0	0.0	**0.6**	0.0	**-0.4**	0.0	0.0	-0.1	-0.1
PPR if 5s 0d	0.0	0.0	0.0	-0.1	-0.1	0.0	**-0.7**	0.0	0.0	0.0	**0.2**	-0.1	0.1
PPR if 4s 1d	0.1	0.0	0.0	-0.1	-0.1	0.0	**-1.3**	0.0	**-0.5**	0.0	0.3	0.1	0.1
PPR if 3s 2d	0.0	0.0	0.0	**-0.1**	0.0	0.0	**-0.2**	0.0	0.1	0.0	0.0	-0.1	0.0
PPR if 2s 3d	0.0	0.0	0.0	0.1	**0.4**	0.0	**-0.3**	0.0	0.2	0.0	0.1	0.0	0.1
PPR if 1s 4d	0.1	0.0	0.0	0.0	0.2	0.0	**1.0**	0.0	**-0.3**	0.0	-0.2	0.1	-0.1
PPR if 0s 5d	0.0	0.0	0.0	0.0	0.0	0.0	**-0.3**	0.1	**-0.4**	0.0	-0.1	-0.3	0.0
Fertility diff across women's edu	**-0.4**	0.0	-0.1	0.0	0.0	0.0	0.2	0.1	**1.0**	0.0	0.0	0.2	0.0
Married men with high edu (%)	0.1	-0.1	0.0	0.0	0.0	0.0	0.2	0.0	**-0.9**	0.0	0.1	**0.5**	0.0
Married women with high edu (%)	-0.1	0.0	0.0	-0.1	0.0	0.0	-0.1	-0.1	**0.5**	0.0	0.0	**0.7**	0.0
Education expenditure per son	0.1	0.0	0.0	-0.1	0.1	-0.1	-0.2	-0.1	**-1.1**	0.1	-0.2	**-0.4**	0.0
Education expenditure per daughter	0.0	0.1	0.1	-0.1	0.0	0.2	**-1.0**	0.1	**0.3**	0.0	-0.2	-0.2	0.0
Prob that next child is a son if 1s 0d	0.0	0.0	0.0	0.0	0.0	0.0	**-0.6**	0.0	**1.5**	0.0	0.0	0.0	0.0
Prob that next child is a son if 0s 1d	0.2	0.0	0.0	0.0	0.0	-0.1	**1.3**	0.0	**-0.8**	0.0	0.0	-0.1	0.0
Prob that next child is a son if 2s 0d	0.0	0.0	0.0	0.0	0.0	0.0	**-0.6**	0.0	**0.6**	0.0	0.0	0.0	0.0
Prob that next child is a son if 1s 1d	-0.1	-0.1	-0.1	0.0	0.0	-0.1	-0.1	0.1	**-0.5**	0.0	-0.1	**-0.2**	0.0
Prob that next child is a son if 0s 2d	0.2	0.1	0.1	0.0	0.0	0.1	**-0.4**	0.0	**-0.4**	0.0	0.0	0.0	0.0
Prob that next child is a son if 3s 0d	0.0	0.0	0.0	0.0	0.0	0.0	**-0.9**	0.0	**-0.3**	0.0	0.0	-0.1	0.0
Prob that next child is a son if 2s 1d	0.0	**0.0**	0.0	0.0	0.0	0.0	0.0	0.0	**0.4**	0.0	0.0	0.0	0.0
Prob that next child is a son if 1s 2d	-0.1	-0.1	-0.1	0.0	0.0	0.0	**0.4**	-0.1	**4.6**	0.0	0.0	0.1	0.0
Prob that next child is a son if 0s 3d	0.1	-0.1	0.0	0.0	0.0	0.0	**1.4**	0.1	**-0.3**	0.0	0.2	0.0	0.1
Prob that next child is a son if 4s 0d	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**-0.3**	0.0	0.0	**0.0**	0.0
Prob that next child is a son if 3s 1d	0.1	0.0	0.0	0.0	0.0	0.0	**-0.2**	0.0	**-2.4**	0.0	0.0	0.0	0.0
Prob that next child is a son if 2s 2d	-0.2	-0.1	-0.1	0.0	0.0	-0.1	0.0	0.1	**-0.6**	0.0	-0.1	**-0.2**	0.0
Prob that next child is a son if 1s 3d	**0.2**	0.0	0.0	0.0	0.0	-0.1	0.1	-0.1	**-0.9**	0.0	-0.1	0.0	0.0
Prob that next child is a son if 0s 4d	0.2	0.1	0.1	0.0	0.0	0.0	**-0.6**	0.1	**-1.2**	-0.1	-0.1	-0.1	0.1
Prob that next child is a son if 5s 0d	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Prob that next child is a son if 4s 1d	-0.1	-0.1	-0.1	-0.1	0.1	-0.2	**0.4**	0.0	**-0.5**	0.0	-0.1	-0.1	0.0
Prob that next child is a son if 3s 2d	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**-0.1**	**0.0**	0.0	0.0	0.0
Prob that next child is a son if 2s 3d	0.1	0.1	0.1	0.0	0.0	0.0	**-0.2**	0.0	**3.7**	0.0	0.0	-0.1	0.0
Prob that next child is a son if 1s 4d	0.0	0.0	0.0	0.0	0.0	0.0	**0.3**	0.1	**-0.5**	0.1	0.0	-0.2	0.0
Prob that next child is a son if 0s 5d	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	**0.2**	**-0.4**	0.0	0.0	-0.1	-0.1

Notes. Numbers in bold blue show the largest elasticities for each moment.

C Estimate the Model to Match Punjab Data

This appendix shows how the model is mapped to the data in Punjab. First, I use the same parameters chosen outside the model. Recall that they are from the literature or set to their data counterparts. Although it is quite likely that dowry as a fraction of household income (δ) and old-age support responsibilities $\{t_{m1}, \kappa_m, t_{f1}, \kappa_f\}$ are different for Punjab from an average Indian state, I do not have enough observations in the data to compute these values for Punjab. In addition, labor market discrimination λ is taken from Deshpande et al. (2018), so its value is kept as well.

The remaining 13 parameters are estimated within the model to match the following 26 data moments in Punjab.

(i) Household savings rate when middle-aged. Source: Indian HDS 2011-2012. [1 target]

(ii) Women's working hours when young and middle-aged as a fraction of their time endowment. Source: Indian HDS 2011-2012. [2 targets]

(iii) PPR conditional on the sex composition of the first one, two, and three children. Source: Indian DHS 2015-2016. [9 targets]

(iv) Fertility difference between women with and without secondary education. Source: Indian DHS 2015-2016. [1 target]

(v) Share of men and women with secondary education. Source: Indian DHS 2015-2016. [2 targets]

(vi) Education expenditure per son and per daughter as a fraction of household income. Source: Indian HDS 2011-2012. [2 targets]

(vii) Probability that the next child is a son conditional on the sex composition of the first one, two, and three children. Source: Indian DHS 2015-2016. [9 targets]

Note that we have fewer moments for (iii) and (vii). Since the sample size for Punjab is small, the parity progression ratios and the probability that the next child is a son conditional on the sex composition of the first four and five children cannot be computed precisely. Therefore, they are dropped from the set of targets.

The estimated parameters are shown in Table A5. The weight of daughters relative to sons (ω) is smaller than in the baseline estimation, suggesting that parents in Punjab have a stronger intrinsic preference for sons. Moreover, the values of μ_ξ and σ_ξ suggest that parents in Punjab incur a lower utility cost from sex selection.

Table A5: Estimated Parameters for Punjab

	Description	Benchmark	Punjab	Target
β	Annual discount factor	0.9765	0.9770	Moments (i)
ν_1	Weight of non-market goods when young	0.2274	0.3158	Moments (ii)
ν_2	Weight of non-market goods when middle-aged	0.5571	0.6763	Moments (ii)
μ_α, σ_α	Distribution of weight of children	-1.1239, 0.5402	-1.0923, 0.3834	Moments (iii)-(vii)
ω	Relative importance of daughters	0.9237	0.9173	Moments (iii)-(vii)
ϵ	Substitution btw sons and daughters	19.93	25.32	Moments (iii)-(vii)
μ_ξ, σ_ξ	Distribution of disutility from sex selection	-3.5493, 0.4464	-3.8135, 0.2791	Moments (iii)-(vii)
η	Importance of education investment in utility	0.5795	0.6017	Moments (iii)-(vii)
e_0	Public education expenditure	0.0311	0.0197	Moments (iii)-(vii)
z	Efficiency of human capital production	12.13	11.74	Moments (iii)-(vii)
τ	Fixed time cost of each child	0.1066	0.0979	Moments (iii)-(vii)

Notes. For Punjab, 26 moments are used to estimate 13 parameters.

Moments (i), (ii), (iv), (v), and (vi) in the data and model are shown in Table A6. Moments (iii) and (vii) are shown in Figures A7 and A8, respectively.

Table A6: Data and Model: Selected Moments for Punjab

Variable	Data	Model
Targeted moments		
Savings rate when middle-aged	0.367	0.368
Female labor supply when young	0.180	0.192
Female labor supply when middle-aged	0.167	0.159
Fertility difference across women's education	0.69	0.66
Share of married men with high education (%)	75.3	100.0
Share of married women with high education (%)	65.5	65.1
Education expenditure per son	0.0664	0.0770
Education expenditure per daughter	0.0575	0.0540
Aggregate moments		
Fertility	2.65	2.53
Sex ratio at birth	1.334	1.327

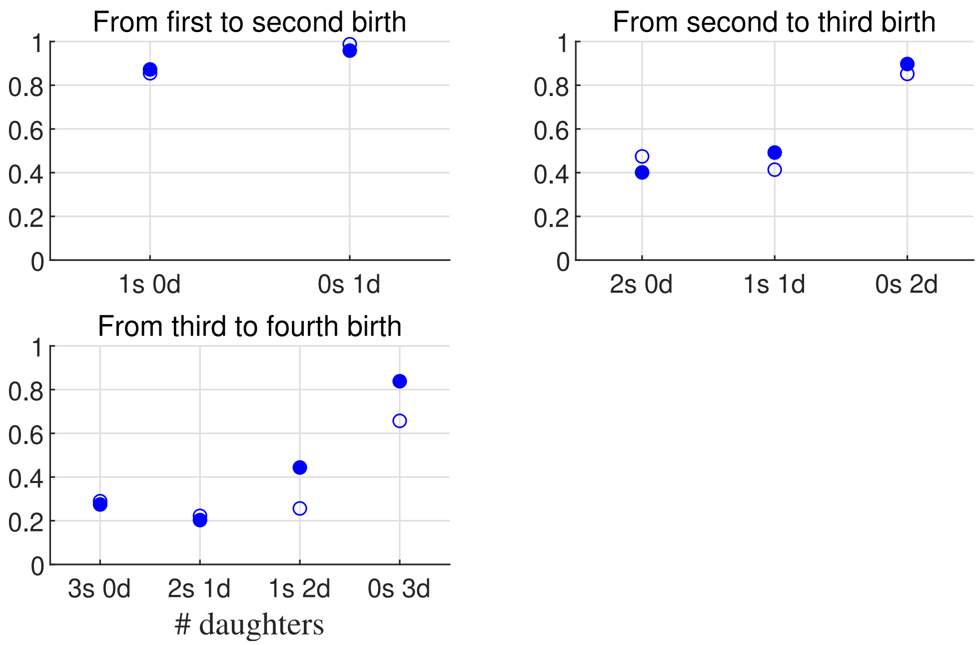


Figure A7. Parity Progression Ratios Conditional on the Sex Composition of Previous Children: Data and Model for Punjab.

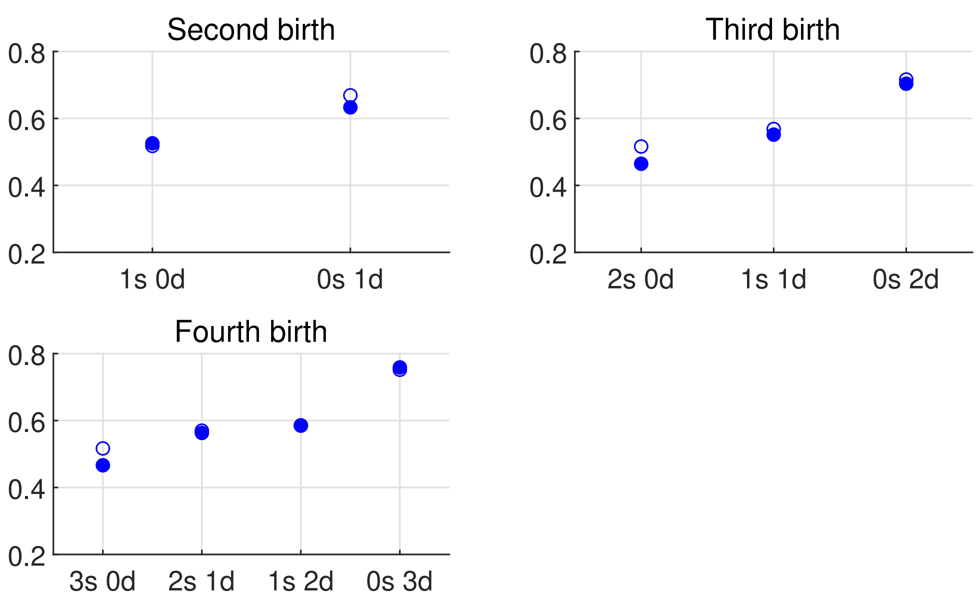


Figure A8. Probability of a Son Conditional on the Sex Composition of Previous Children: Data and Model for Punjab.

D Additional Results

D.1 Parity Progression and Sex Selection in Some Counterfactual Experiments.

Figure A9 shows the parity progression ratios conditional on the sex composition of existing children if the gender differences in the economic factors are removed. Parents are less likely to practice a son-biased fertility stopping rule.

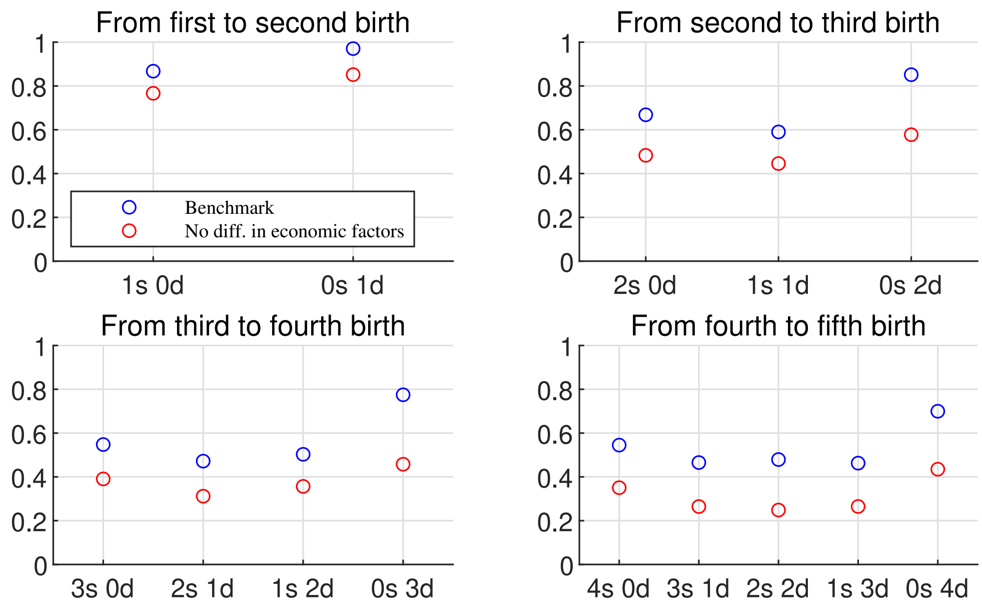


Figure A9. Parity Progression Ratio Conditional on the Sex Composition of Previous Children if No Gender Difference in Economic Factors.

Figure A10 shows the probability that the next birth is a son in this case. Compared with the benchmark economy, where parents aborted female fetuses disproportionally if they have few sons, parents in this counterfactual economy almost do not practice son-biased sex selection even if they have no son.

Figure A11 shows the parity progression ratios conditional on the sex composition of existing children if both the gender differences in the economic factors and intrinsic sex preference are removed. Parents do not practice a son-biased fertility stopping rule.

Figure A12 shows the probability that the next birth is a son in this case. Compared with the benchmark economy, where parents aborted female fetuses disproportionally if they have few sons, parents in this economy do not practice son-biased sex selections.

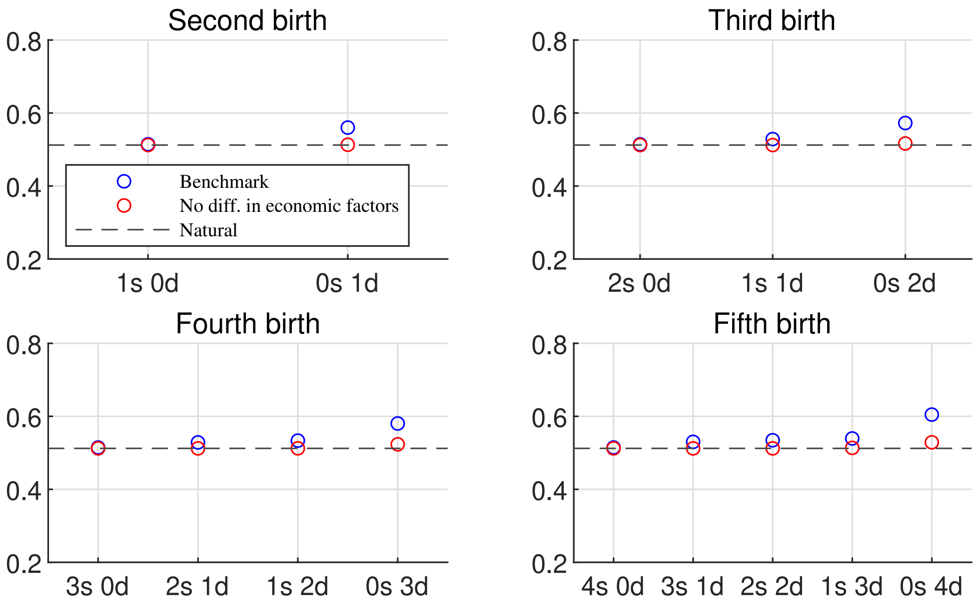


Figure A10. Probability of Having a Son Conditional on the Sex Composition of Previous Children if No Gender Difference in Economic Factors.

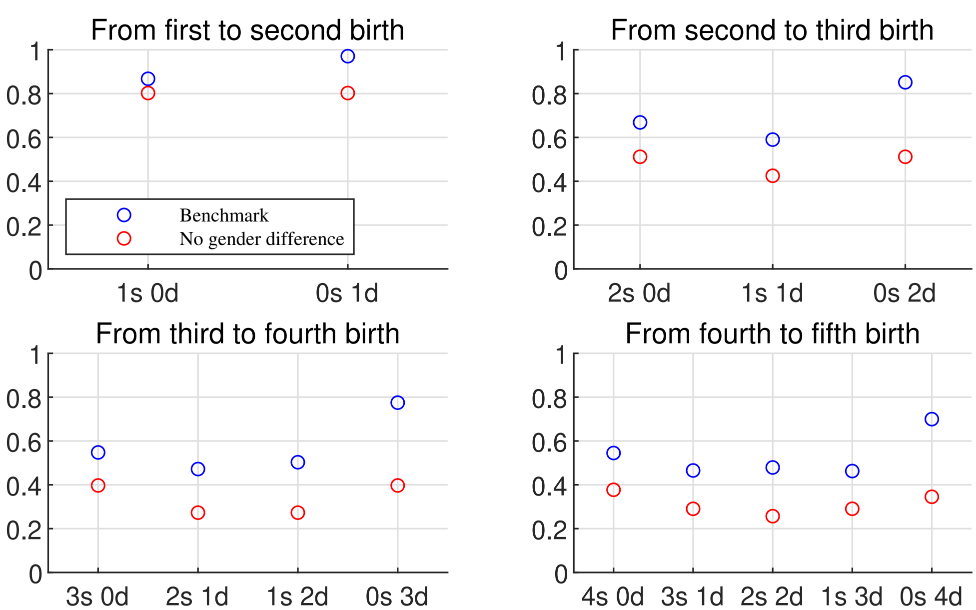


Figure A11. Parity Progression Ratio Conditional on the Sex Composition of Previous Children if No Gender Difference.

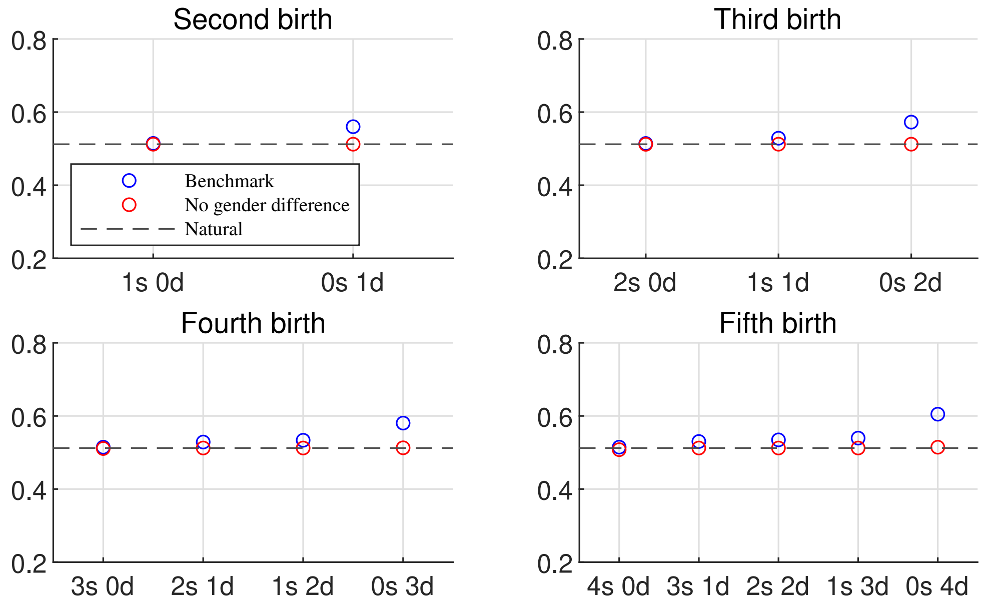


Figure A12. Probability of Having a Son Conditional on the Sex Composition of Previous Children if No Gender Difference.

D.2 Other Subsidy Rates

Table A7 shows the effects of the subsidy for female births for different levels of subsidies. As the subsidy goes up, the sex ratio at birth decreases. Since the fixed cost of births is reduced, fertility increases while education expenditure decreases for both boys and girls. This leads to a smaller share of men and women with high education, with the gender gap largely unchanged. In addition, a higher fertility rate decreases young women's labor supply.

Table A8 shows the effects of the subsidy for female education for different subsidy rates. As the subsidy rate goes up, the sex ratio at birth decrease monotonically. Moreover, there is a monotonic increase in education expenditure for daughters, not only because the subsidy directly incentivizes parents to invest more in daughters' education but also because parents shift resources from sons to daughters. Education expenditure for boys, however, may increase or decrease since there are two competing effects. On the one hand, a higher subsidy rate decreases fertility and increases education expenditure for boys. On the other hand, a higher subsidy rate for female education motivates parents to shift resources from sons to daughters. Overall, as the subsidy rate goes up, the gender education gap firstly closes and then gets reversed. In addition, with a higher subsidy rate, women supply more labor to the market.

Table A7: Effects of the Subsidy for Female Births

		Subsidy		
	Benchmark	0.01	0.02	0.03
Sex ratio at birth	1.139	1.086	1.063	1.054
Fertility	3.07	3.28	3.55	3.84
Education expenditure per son	0.0570	0.0528	0.0478	0.0427
Education expenditure per daughter	0.0379	0.0343	0.0303	0.0265
Men with high education (%)	64.6	59.9	54.0	47.8
Women with high education (%)	49.2	44.1	38.5	32.8
Gender gap in education	15.4	15.8	15.5	14.9
Female labor (young)	0.250	0.232	0.207	0.178
Female labor (middle-aged)	0.270	0.278	0.284	0.291
Labor income tax rate	0	0.006	0.014	0.024

Table A8: Effects of the Subsidy for Female Education

		Subsidy rate		
	Benchmark	10%	20%	30%
Sex ratio at birth	1.139	1.113	1.091	1.071
Fertility	3.07	2.93	2.82	2.74
Education expenditure per son	0.0570	0.0572	0.0556	0.0524
Education expenditure per daughter	0.0379	0.0472	0.0568	0.0669
Men with high education (%)	64.6	65.5	65.0	63.1
Women with high education (%)	49.2	61.5	72.7	82.6
Gender gap in education	15.4	4.1	-7.7	-19.5
Female labor (young)	0.250	0.280	0.305	0.326
Female labor (middle-aged)	0.270	0.293	0.315	0.334
Labor income tax rate	0	0.004	0.010	0.020

D.3 Other Replacement Rates of Pension

Table A9 presents the effects of pensions with different replacement rates. As the replacement rate goes up, the sex ratio at birth decreases slowly. There is also a decrease in education investment for both boys and daughters, so the share of men and women with high education decreases. The gender education gap does not change much. In addition, women supply less labor. Fertility is negatively affected, but the effect is very small.

Table A9: Effects of the Pension System

		Replacement rate		
	Benchmark	30%	50%	70%
Sex ratio at birth	1.139	1.103	1.092	1.084
Fertility	3.07	3.05	3.03	3.01
Education expenditure per son	0.0570	0.0542	0.0540	0.0538
Education expenditure per daughter	0.0379	0.0389	0.0384	0.0381
Men with high education (%)	64.6	59.1	56.6	54.4
Women with high education (%)	49.2	46.9	44.4	42.2
Gender gap in education	15.4	12.2	12.2	12.2
Female labor (young)	0.250	0.279	0.295	0.312
Female labor (middle-aged)	0.270	0.283	0.279	0.274
Savings rate (middle-aged)	0.361	0.324	0.287	0.245
Labor income tax rate (men)	0	0.085	0.141	0.198