Missing Women: A Quantitative Analysis*

Yongkun Yin

CEMFI

Job Market Paper

November 19, 2022

Please click here for latest version

The sex ratio, males per female, is well above one in China, India, and other South and East Asian countries. Parents in these countries want more boys, exercise sex-selective abortions, and invest less in their daughters' education. Why do parents favor sons over daughters? What policies can be effective in normalizing the sex ratio? To answer these questions, I build an overlapping-generation model of fertility, sex selection, the quantity-quality trade-off, and marriage and estimate it 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 these economic factors are removed, the sex ratio at birth (SRB) would reduce from 1.14 to 1.05. The fertility rate would drop from 3.1 to 2.4, and the share of women with secondary education would increase from 49% to 72%. The sons would also benefit from lower fertility, and the share of men with secondary education would rise from 65% to 79%. Once the economic factors become gender-neutral, eliminating intrinsic son preferences has a small additional effect. A subsidy for female births or female education, commonly-implemented policies in India and elsewhere, can both reduce the SRB. However, the former increases fertility and reduces children's education and women's labor supply, while the latter has the opposite effects. Finally, a pay-as-you-go pension system can lower the SRB to 1.09, but it also reduces children's educational attainment, as parents value them now less.

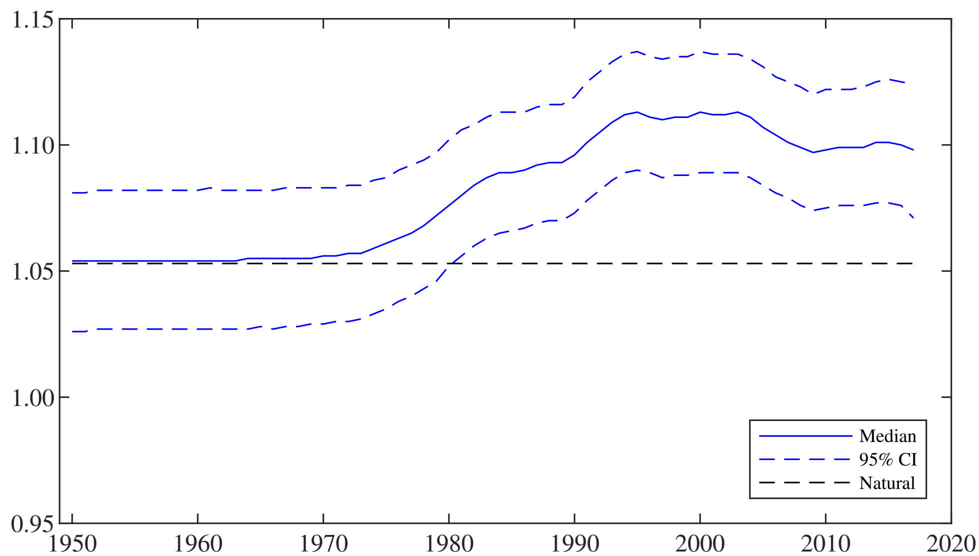
Keywords: Sex ratio; Fertility; Education; Quantity-quality trade-off; India.

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, Tom Zohar, and seminar participants at CEMFI and Shandong University for insightful comments and discussions. I acknowledge the funding from Fundación Ramón Areces.

1 Introduction

In several South and East Asian 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 fairly constant across time and space due to biological reasons (Orzack et al., 2015). The SRB in India began to increase in the 1980s, when ultrasound machines, which can be used for prenatal sex detection, were increasingly available. In the 1990s, the ratio exceeded 110 male births per 100 female births, well above the natural level.



Notes. Data are from Chao et al. (2019).

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

A biased sex ratio at birth or son preference, as it is often referred to in the literature, 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, excessive female child mortality, and low education levels of women (see, among others, Hesketh and Zhu, 2006; Dyson, 2012; Barcellos et al., 2014).

[1] Most studies focus on developing countries. However, there is also evidence 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).

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 dowry 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 the parental utility function, i.e., an intrinsic son preference.[2] Despite these insights, surprisingly little is known about the relative importance of these factors, how they interact and affect the aggregate economy, and what policies can be most effective in correcting the biased sex ratio and its associated problems. This paper tries to fill this gap.

Son preference can have important effects on the level of fertility and children's education. Since parents tend to have another birth if they do not have sons (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 desired number of sons, parents 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, the availability of sex selection reduces uncertainty about the sex of new births, increasing parents' incentive to have more. Second, sex selection technologies can change the trade-off between prenatal sex selection and postnatal sex discrimination. Since girls

[2]There are a few reasons for this 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 assuming 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.

born despite the accessibility of the sex selection technologies are more "wanted", 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 do these factors affect the level of fertility, women's labor supply, and household income in equilibrium? What are the impacts of sex selection technologies? What policies can normalize the sex ratio, and how can the policies affect fertility, female labor supply, and household income?

To answer these questions, I build a quantitative overlapping-generation (OLG) model embedding the standard quantity-quality trade-off and marriage, which incorporates various causes of missing women and a sex selection technology. Parents pay dowries to their married daughters and receive transfers from their adult children when they get old. However, they mainly rely on their sons for old-age support. There is also labor market discrimination, which acts as a wedge between women's productivity and wage. These economic factors affect the 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 potentially value daughters less than sons. Fertility decisions are made sequentially conditional on the sex composition of the previous children. In their sequential fertility decisions, parents can use a sex selection technology to ensure a male or female birth at a monetary and a utility cost. Parents' fertility, sex selection, and education decisions are forward-looking. Furthermore, the model has rich interactions among households through marriage, dowries, and old-age support.

The model parameters are calibrated to match the Indian data. The model can replicate the moments that are commonly used to discipline OLG models of fertility and investment in children, such as households' savings rate, women's working hours, women's education gradient in fertility, education expenditure per son and daughter, and the educational attainment of men and women's education. Furthermore, it can reproduce the observed probabilities of having another birth and the likelihood that the next birth will be a boy, conditional on the sex composition of the previous children. These moments are critical to identifying how substitutable sons and daughters are in the parents' preferences and how strong the intrinsic son preferences are.

Given the estimated parameters, I first compare my model's outcomes with empirical papers that study the effects of government programs on missing women. I focus on a program called *Dhanlakshmi* that subsidized female births and female education in seven Indian states between 2008 and 2013. I show that conditional cash transfers of the magnitudes comparable to this program in the model increase the probability of a female birth, and the model's short-run responses are in line with evidence provided by a recent working paper by Biswas et al. (2021). I also explore 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% and reduces the share of women completing secondary education by 1 pp. This finding is consistent with the available evidence on the impact of sex selection technology (Nandi and Deolalikar, 2013; Rastogi and Sharma, 2022).

Next, I run experiments to quantify the relative contribution of different factors to missing women and their effects on fertility, education, female labor supply, and household income. I start by removing the gender difference in one economic factor at a time and then remove all of them simultaneously. Finally, 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 1.14. 65% of men complete secondary education, but only 49% of women achieve this. If there is no gender difference in the economic factors, the SRB will drop to 1.05. The share of men with secondary education would rise to 79%, while the share of women would rise more significantly to 72%. If the intrinsic sex preference is also eliminated, the SRB would barely drop further. The share of men with secondary education would fall slightly to 78%, but the share of women with secondary education would rise slightly to 77%. Hence, parents' incentives to have sons impact their quality-quantity trade-off, and parents end up with too many children. This results in lower investment in boys and girls, but the impact on girls is much more substantial.

The second finding is that economic factors have a much more significant effect than intrinsic son preference on other outcomes of interest, including fertility and female labor supply. In the benchmark model, the fertility rate is 3.1, and women supply 26% of their time to the market. If there is no gender difference in the economic factors, the fertility rate will drop by 0.7 to 2.4, women would supply 45% of their time to the market, and the average household income would increase by 25%. If intrinsic son preference is further eliminated, the marginal effects are minimal.

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 would increase to 1.23 instead of decreasing. The reasons are twofold. First, the value of daughters does not increase unless their income is used to provide old-age support, while 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 sons if they do not abort female fetuses. Both factors interact and generate incentives for parents to select a male birth.

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 are widely used by the state governments of India. 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 in different directions. By lowering the fixed cost of each birth, a subsidy for female births increases fertility, which reduces education expenditure per child and decreases women's labor supply. Indeed, a subsidy of 16% of the annual income of a male without secondary education would raise the fertility rate to 3.6, reduce the share of men and women with secondary education to 54% and 39%, respectively, reduce women's labor supply to 25% of their time, and reduce household income by 5%. In contrast, by motivating parents to invest more in girls' education, a female education subsidy reduces fertility and increases women's labor supply. Indeed, a subsidy rate of 19% would reduce the fertility rate to 2.9, increase the share of men and women with secondary education to 65% and 57%, respectively, increase women's labor supply to 31% of their time, and increase household income by 2%.

Next, I assess the impacts of a pension system. As another source of income after retirement, a pension weakens the role of old-age support provided by children and the gender difference in old-age support. Therefore, fertility and education expenditures would be reduced at the same time as a reduction in the sex ratio. More specifically, I consider a pay-as-you-go pension system with a replacement rate of 50%, which is the rate targeted by the Employees Provident Fund Scheme, the largest pension scheme in India. The SRB would drop to 1.09, fertility would drop slightly, and the share of men and women with secondary education would drop to 57% and 45%, respectively.

This 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 other 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. This paper complements these studies by investigating the role of sex selection technologies and emphasizing the effect on the sex ratio.

This paper is also related to the economic and demographic literature on the causes of a 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 focus on one factor, my study uses a quantitative model to quantify the relative importance of these factors.

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 and short-run effects of a small-scale policy. A quantitative model allows me to assess the impact at the aggregate scale in the long run.

The remainder of the paper is organized as follows. Section 2 documents some empirical facts that motivate this study. Section 3 describes the model. Section 4 calibrates the model. Section 5 assesses the causes and consequences of missing women. Section 6 examines possible policies to reduce the imbalanced sex ratio 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 levels.

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, 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

[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.

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 themselves and their husbands. 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.

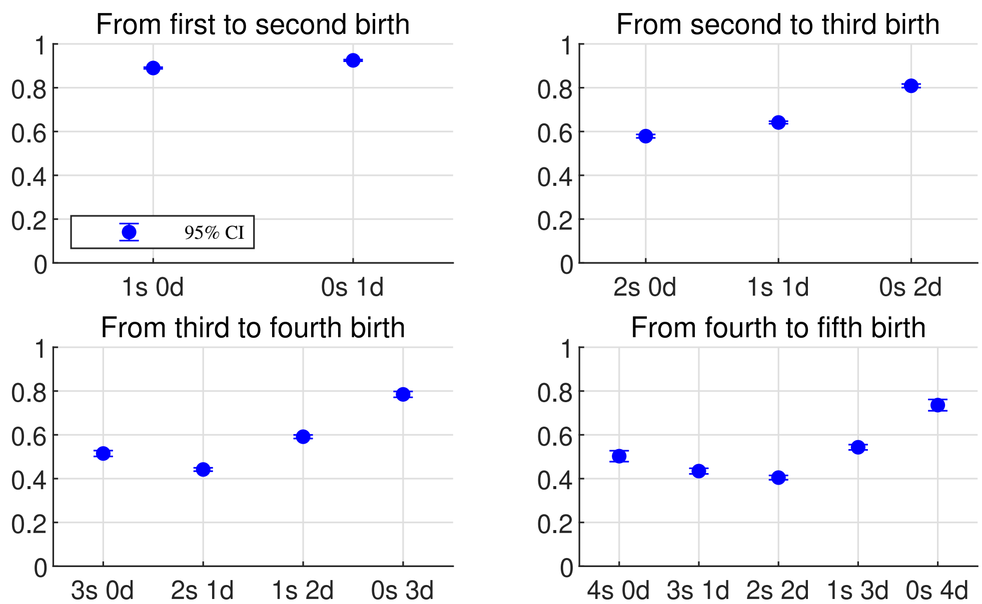
The Indian DHS data in 2015-2016 suggest that about 19% of parents want more sons than daughters but only 3% want the opposite, suggesting a strong favor for sons.[5] What will parents do if they want sons but do not have (enough) sons? A typical practice is to continue childbearing until reaching their ideal number of sons (e.g., Clark, 2000; Altindag, 2016). This subsection examines how the parental decision on whether to have another child depends on the sex composition 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 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. 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 sons and two daughters. The asymmetry in the PPR suggests that parents have a preference for sons and want more sons than daughters.

Second, parents are more likely to have another birth if their children have an imbal-

[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.

[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).



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.

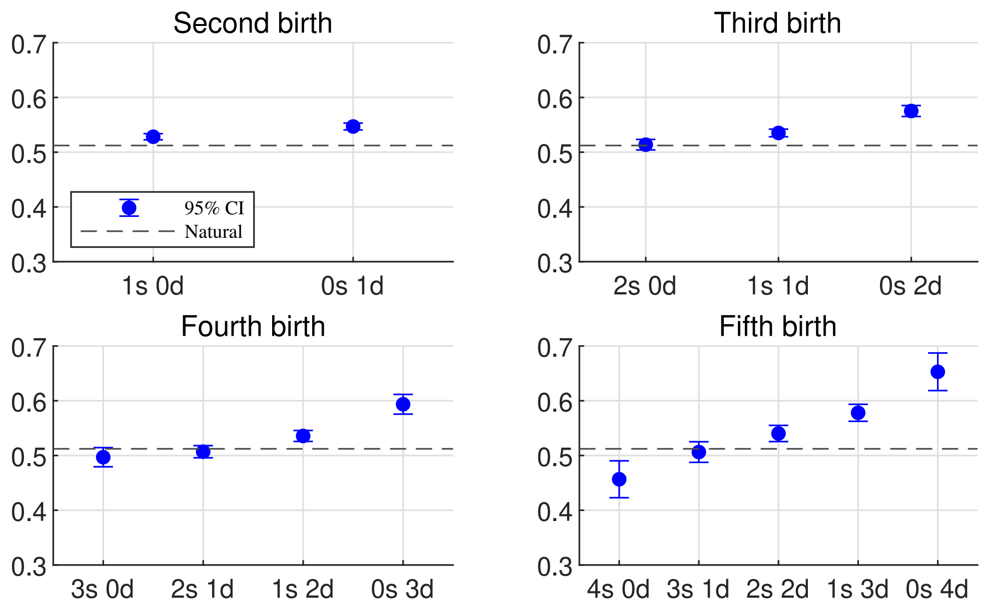
anced sex composition, i.e., if they either have too few sons or too few daughters. The pattern is most salient among women with three or four children.[7] Consider women with four children. If they have two sons and two daughters, only 40.5% of them will have another birth. However, if they have only daughters or only sons, the ratio increases sharply to 50.3% and 73.6%, respectively. The U-shaped PPR suggests that parents do not treat sons and daughters as perfect substitutes and that they want a balanced sex composition of children.

Fact 2. Son-biased Sex Selection.

In the mid-1980s, prenatal sex detection became 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 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

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

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 Existing Children.

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.

In addition, it is noteworthy that parents may 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.

Finally, besides sex-selective abortion, female children may suffer a high mortality rate due to child neglect, especially when parents have no access to ultrasound machines (Das Gupta, 1987; Arnold et al., 1998). While the mortality rate is higher for male children than female children in most parts of the world (Pongou, 2013), the Indian DHS data in 2015-2016 suggest that the under-five mortality rate is about 59 per 1000 live births for female children, which is even larger than the mortality rate for male children (57 per 1000 live births).

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. I run ordinary least square (OLS) regressions 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. 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 to boys rather than girls. As will be shown next, girls receive less education than boys.

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. *** $p < 0.01$.

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 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 reveals 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 Gradient in 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. I run an OLS regression 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.

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. *** $p < 0.01$.

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 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 age and education 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 show that similar patterns emerge for women with and without secondary education. Furthermore, the findings are robust to using previous survey waves in 1992-1993, 1998-1999, and 2005-2006. I also show similar parity progression ratios in Nepal and Vietnam, two countries where the survey evidence indicates that women want more sons than daughters (Fuse, 2010). Similarly, Fact 2, the sex selection conditional on the sex composition of existing children, is observed for women with and without secondary education. In the earlier waves of the data, particularly for the 1992-1993 and 1998-1999 waves, parents were less likely to select their children's sex, which reflects the unavailability of sex-selection technologies for older cohorts. Sex selection can also be detected in Nepal and Vietnam.

Regarding Facts 3 (the quality-quantity trade-off) and 4 (gender education gap), I show that the results do not change with samples of men and women from different age groups (ages 23-32, 33-42, and 43-52). Finally, for Fact 5 (women's education gradient in fertility), I show that the results are robust to using different waves of the Indian DHS (in 1992-1993, 1998-1999, and 2005-2006) and data from Nepal and Vietnam.

3 The Model

In this section, I build an overlapping-generation model of fertility, sex selection, 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 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 bride's parents. 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.[8] Parents can select their children's sex, at a monetary and utility cost.

Middle-aged adults work, consume and save for their old age.[9] 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 sex composition. Old adults do not work and simply consume their savings and transfers from their children. Next, I describe the model in detail.

[8]Hirway and Jose (2011) document the household division of labor in India based on the 1998-1999 Time Use Survey. They show that on average, in a week a woman spends 30.0 hours on market work, 6.1 hours on the collection of fuel wood, fodder, water, and other raw materials, 28.9 hours on household maintenance, and 4.5 hours on care of the children and other adults in the family. A man spends 54.6 hours on market work but only 4.0 hours on the other three activities.

[9]The assumption that only middle-aged parents can save is done for computational reasons.

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). Women are discriminated against in the labor market, and there is a wedge between their wage rate and their productivity. I assume that the discrimination costs are redistributed in lump-sum back to working households, and each household receives a transfer T.[10] Upon marriage, the young couple receives a dowry (d) from the wife's parents. The household income is,

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

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

Starting with 0 children, a young couple decides whether to give 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. However, they need to pay a monetary cost π and incur a utility cost ξ. The monetary cost captures what a 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, and actual side effects that reduce people's well-being, which can be different across individuals.[11] Therefore, I assume that ξ is drawn randomly from a log-normal distribution after marriage, i.e., $\log(\xi) \sim N(\mu_\xi, \sigma_\xi)$.

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

[10]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.

[11]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)). The cost differences might also reflect the additional difficulty that people in more remote locations might face accessing sex selection technology.

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 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.[12] 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;

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

Delventhal et al., 2021).[13]

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 $+\infty$.[14]

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.[15,16] 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 par-

[13] 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.

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

[15] 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 the dowry is proportional to the father's income.

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

ents. 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 is a function of the number of brothers and sisters of the husband, or the number of male and female children in his natal family, $\{q_m^{p,m}, q_f^{p,m}\}$. Similarly, t_f is a function of the number of male and female children in the wife's natal family, $\{q_m^{p,f}, q_f^{p,f}\}$. Therefore, the consumption of market goods by the middle-aged couple is,

$$c_2 = [1 - t_m(q_m^{p,m}, q_f^{p,m}) - t_f(q_m^{p,f}, q_f^{p,f})]I_2. \tag{7}$$

Old Married Households

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

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

[17] The data from the Longitudinal Ageing Study in India in 2017-2018 suggest that less than 5% of

children's income is transferred to parents. In the section on 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 in old age is the savings from middle age and the transfers from children, i.e.,

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

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 of 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 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

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. 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).

children, who are now 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.

For a middle-aged household, the state variables is $\mathbf{x}_2 = \{h_m, h_f, q_m, q_f, \mathbf{x}_1^{k,m}, \mathbf{x}_1^{k,f}, \mathbf{x}_3^{p,m}, \mathbf{x}_3^{p,f}\}$, where $\mathbf{x}_1^{k,m}$ and $\mathbf{x}_1^{k,f}$ are the state of the male and female children, who are young, and $\mathbf{x}_3^{p,m}$ 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 are given by $\mathbf{x}_1 = \{h_m, h_f, q_m, q_f, q_s, \mathbf{x}_2^{p,f}, \alpha, \xi\}$, 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 at different ages. Note that, in the budget constraint for market goods, $d(\mathbf{x}_2^{p,f})$ is the dowry payments from the 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 exposition, 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 the 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 $\bar{q} - 2$, $\bar{q} - 3$, $\ldots$, 0 in the same way, where $\bar{q}$ is the maximum possible number of children

3.3 Household Formation

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 men with human capital h by $\tilde{Q}_m^h$, then we have

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

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

3.4 Equilibrium

I focus on a 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 decide on female labor supply and savings. Furthermore, given these decisions and the household formation rule, the distribution of individuals by educational attainment and the distribution of households of different ages across their states are constant.

To find a stationary equilibrium, I solve a computational fixed point problem. 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 sons and daughters conditional on their household types. Given this initial guess, household problems are solved, and household formation takes place. The household decisions allow me to update the initial guess, and the process continues until all the distributions and decisions converge.[18]

4 Quantitative Analysis

In this section, I discuss the calibration procedures and the model fit. These parameters are reported in Table 3. 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. The average years of schooling in the sample is 6.9 years. 34.3% of males and 46.7% of females have low human capital, while the rest have high human capital.

[18] Although it is impossible to establish the uniqueness of the fixed point theoretically, the solution algorithm converges to the same stationary equilibrium from several different initial guesses.

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 real interest rate in India in 1978-2020.[19]

To compute human capital levels for 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 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*.[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%

[19] This is based on World Bank data. See https://data.worldbank.org/indicator/FR.INR.RINR?locations=IN.

[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.

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 26% (19% with a low-human capital wife and a high-human capital husband and 7% with a high-human capital wife and a low-human capital husband). Since in the data 53.3% of females and 65.7% of males are high-skilled for this age group, I choose ψ so that given the household formation protocol, the share of different marriages matches the data. Hence, if the benchmark economy matches the educational attainment of males and males, given ψ, it will generate the right degree of assortative mating.

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 process until they conceive a child 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 in the 1990s (about 10-20 dollars at the exchange rate at that time) (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 $750\times2+1661\times1 = 3111$

[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.

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, $\{t_{m1}$, κ_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

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

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 $\{R_{m1}$, κ_m, R_{f1}, $\kappa_f\}$, and, given an estimate for ϕ, to recover $\{t_{m1}$, κ_m, t_{f1}, $\kappa_f\}$. Denote the total time spent with children (time spent with each child times the number of children) by R_c. R_c is given by

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

Hence, as long as parents have sons, they reside with them, and the amount of time they

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

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 4 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 the fraction of children's income transferred to their parents during coresidence, 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. Given that $\kappa_m = 0.972$, for parents with a son, the total transfer will increase by 0.5-1.9% as they have each additional son, and will not change if they have additional daughters. Given that $\kappa_f = 0.925$, for parents with a daughter, the total transfer will increase by only 1.4-5.1% as they have each additional daughter but will increase more substantially by 74% if they have an additional son.

[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 3: 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 the father's annual income	0.6	REDS 1999
π	Monetary cost of sex selection as a fraction of the annual income of men without secondary education	0.1	Arnold et al. (2002), Ganatra and Hirve (2002)
t_{m1}, κ_m	Transfer from sons (with a son)	0.207, 0.972	LASI 2017-2018
t_{f1}, κ_f	Transfer from daughters (without a son)	0.119, 0.925	LASI 2017-2018

Table 4: 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.

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) Parity progression ratios conditional on the sex composition of existing children, as shown by Fact 1 in Section 2. [20 targets]

(ii) Probability that the next child is a son conditional on the sex composition of existing children, as shown by Fact 2 in Section 2. [20 targets]

(iii) Education expenditure per son and per daughter as a fraction of household income. [2 targets]

(iv) Fertility difference between women with and without secondary education. [1 target]

(v) Share of married men and women with secondary education. [2 targets]

(vi) Women's working hours when young and middle-aged as a fraction of their time endowment. [2 targets]

(vii) Household savings rate when middle-aged. [1 target]

Table 5 shows moments (iii), (iv), (v), (vi), and (vii) in the data and model. Moments (i) and (ii) are shown in Figures 4 and 5, respectively. The model can replicate the data well. In particular, it captures the son-biased fertility-stopping rule and the son-biased sex selections. The model predicts a SRB of 1.139, which is somewhat smaller than the data moment (1.184). However, women tend to underreport female children that did not survive, so the moment in the data is probably overestimated (Rosenblum, 2013).[27]

It is clear that in the model, each parameter may affect multiple moments, and each moment may be informative about various parameters. 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). In Table A4, I present the percentage changes in each parameter associated with a one-percent change in each moment, and the discussion here reflects this more formal analysis.

A few parameters can be more easily mapped into particular moments. The discount

[27] 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. However, the overestimation for later births is smaller.

Table 5: Data and Model: Selected Moments

Variable	Data	Model
Targeted moments		
Education expenditure per son (% of household income)	5.05	5.70
Education expenditure per daughter (% of household income)	4.26	3.79
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
Female labor supply when young	0.235	0.250
Female labor supply when middle-aged	0.264	0.270
Savings rate when middle-aged (%)	36.7	36.1
Aggregate moments		
Fertility	3.04	3.07
Sex ratio at birth	1.184	1.139

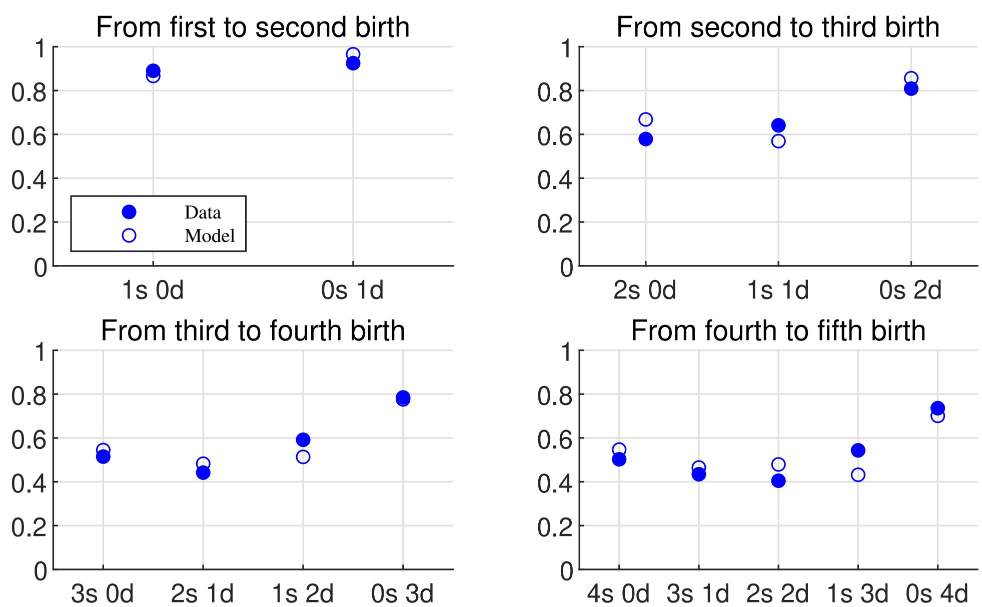


Figure 4. Parity Progression Ratio, Conditional on the Sex Composition of Existing Children (Data and Model).

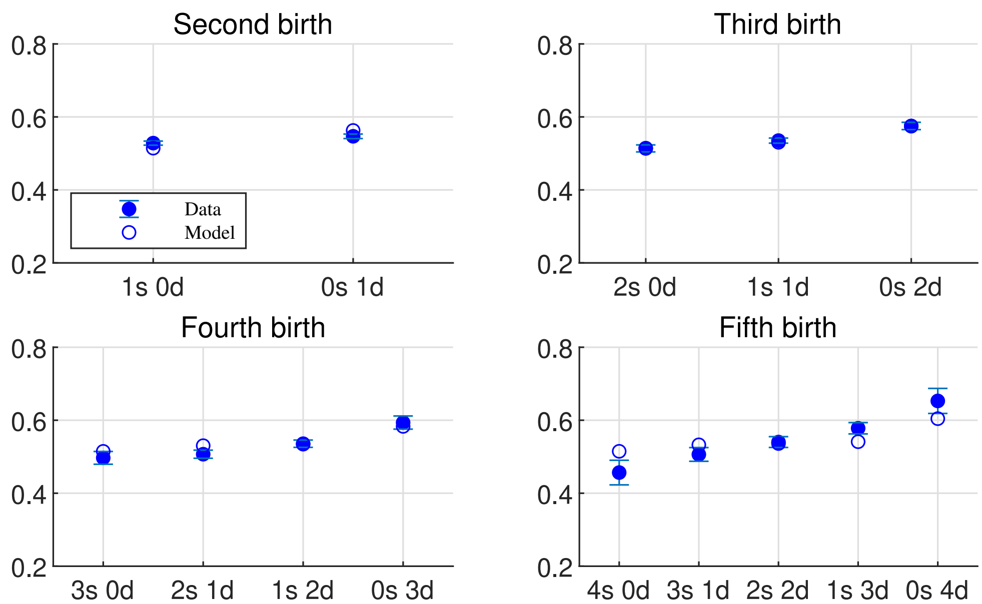


Figure 5. Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children (Data and Model).

factor, for example, mainly determines the savings rate, while ν_1 and ν_2 affect women's labor supply. ν_1 also affects fertility decisions since each birth requires a time commitment from the mother. The efficiency of human capital production (z) determines the educational attainment in the population, while public education expenditure (e_0) affects education expenditure per son and daughter. The relative importance of children (η) and the fixed time cost of each birth (τ) also influence women's fertility rate. At the same time, the relative importance of daughters (ω) also affects education expenditure per son and daughter by altering the parental quantity-quality trade-off.

Finally, the distribution of the utility weight on children $\{\mu_\alpha, \sigma_\alpha\}$, the distribution of disutility from sex selection $\{\mu_\xi, \sigma_\xi\}$, and the elasticity of substitution between sons and daughters (ϵ) are identified by the parity progression ratios (Figure 4) and the sex selection decisions (Figure 5).

The estimated parameters are shown in Table 6. 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. In Figures A8 and A9 in Appendix B, I compare the parity progression ratios and the sex selection decisions when the elasticity of substitution between sons and daughters, ϵ, is 1. When boys and girls are much less substitutable for parents, the PPR takes a symmetric U-shape. The probability of another son (sex selection) now depends much more on their relative scarcity. Hence, sex selection becomes much more strongly correlated with the number of sons that parents have than in the data. 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 indicates that each birth takes about 11% of the mother's time.

Table 6: Estimated Parameters

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

Notes. 48 moments are used to estimate 13 parameters.

4.3 Non-targeted Moments

In this subsection, I show the model fit for some non-targeted moments, i.e., the ones not used in estimation. Figures 6 and 7 show the parity progression ratios (PPR) and the probability that the next child is a son conditional on the sex composition of existing children for women with different education levels. For the PPR (Figure 6), less educated women have higher probabilities of another birth, but the shapes of PPRs do not differ much by education. The model can replicate this feature of the data.

Table 7 presents the fertility rate, the SRB, and education expenditure per son and daughter, again for women with different education levels. Again the model does a good job of capturing the effects of mothers' education on these outcomes. One exception is that the model predicts that women with more education are more likely to select a male birth, while there is no significant difference in the SRB between less and more educated women in the data.[28,29] Finally, Table 8 reports the distribution of married men, married women, and couples based on their education in the model and the data.

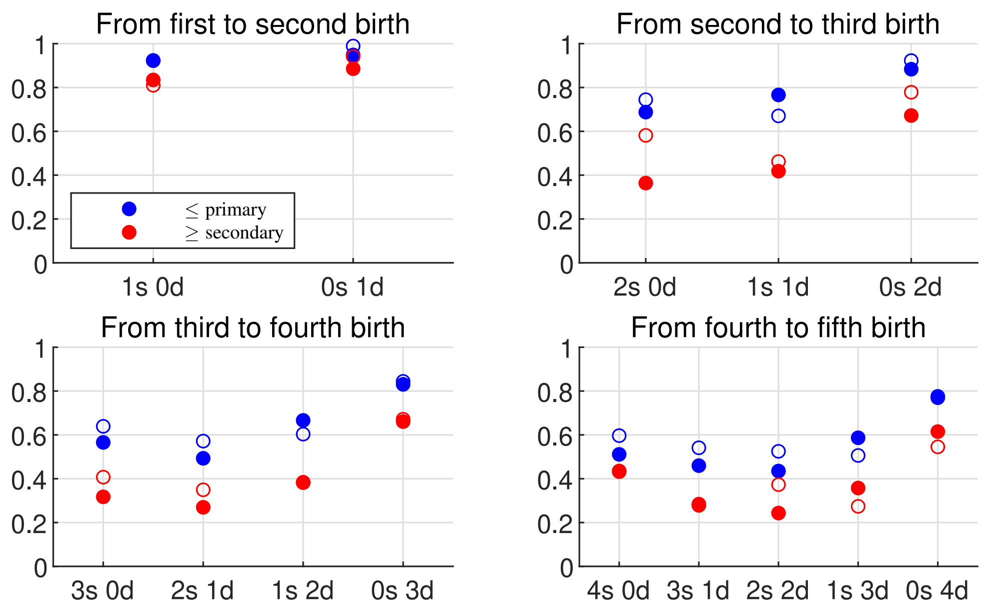


Figure 6. Parity Progression Ratio, Conditional on the Sex Composition of Existing Children by Women's Education (Data and Model).

[28]The relationship between education and sex selection is not a straightforward one, as 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. Furthermore, they have more incentives to select a male birth because they have fewer children and are more likely to end up with no sons if they do not select their children's sex (Jayachandran, 2017). The model captures the latter mechanism but leaves out the former.

[29]As noted before, the SRB in the data is probably overestimated. The bias can 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).

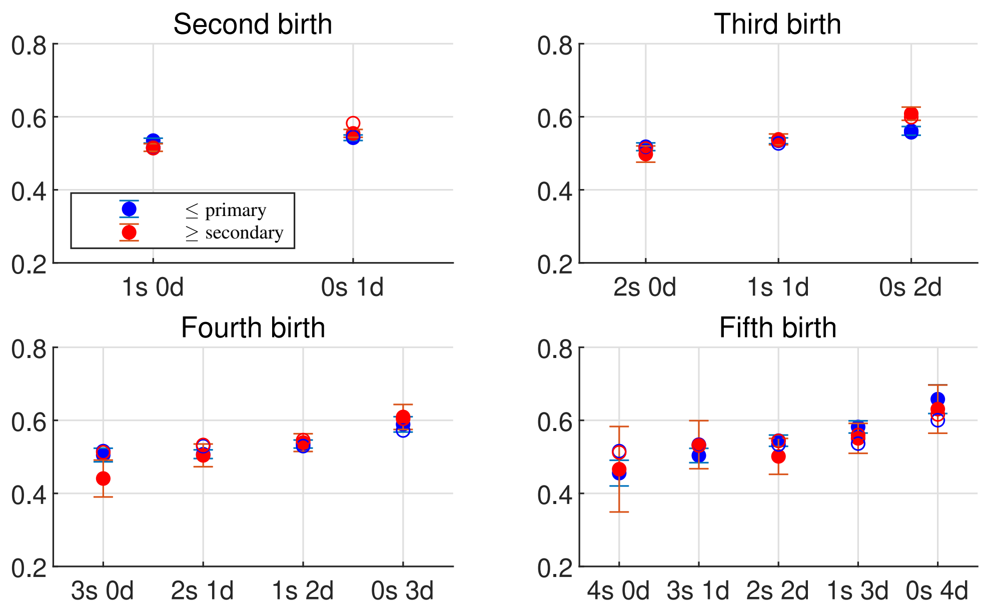


Figure 7. Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children by Women's Education (Data and Model).

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 (% of household income, low)	4.39	4.78
Education expenditure per son (% of household income, high)	6.19	6.65
Education expenditure per daughter (% of household income, low)	3.36	2.85
Education expenditure per daughter (% of household income, high)	5.74	4.75

Table 8: Distribution of Married Men, Women, and Couples (%)

		Men	
		Primary or lower	Secondary or higher
Women	Primary or lower	27 (22)	19 (29)
	Secondary or higher	7 (4)	46 (45)

Notes. Model predictions in parentheses.

4.4 Conditional Cash Transfers: Model vs. Available Estimates

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 predictions that are in line with available estimates.

In 2008, *Dhanlakshmi* was introduced in 11 blocks (the next administrative level below a district) across seven states: Andhra Pradesh, Bihar, Chhattisgarh, Jharkhand, Odisha, Punjab, and Uttar Pradesh.[30] 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 exchange rate at that time) 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

[30] As of August 2022, there are 36 states/union territories and 776 districts in India. On average, there are about 8 blocks in each district.

secondary education in the HDS data in 2011-2012. With sizable benefits, such a program is likely to cause observable changes in fertility and sex selections both in the data and in the model.[31]

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. 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.

To simulate the program in the model, I focus on the transfers for female births and education.[32] 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 when parents are young, while the transfer for female education is received in 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.

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

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

The *Dhanlakshmi* program is first simulated in a short-run partial equilibrium, where the distribution of households by human capital levels and other characteristics, parental expectations about children's marriages and transfers, etc., remain unchanged. Next, it is simulated in a long-run stationary equilibrium, where the household distribution is updated and transfers are 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 short run and 3.6 pp in the long run, 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 short run and 0.131 in the long run, which is in line with the positive effect of the program on fertility risk estimated by Biswas et al. (2021). Overall, the results suggest that the model can deliver reliable predictions in counterfactual analyses.

Table 9: Effects of *Dhanlakshmi*

	Biswas et al. (2021)	Model (short run)	Model (long run)
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		2.53	2.53
Increase in fertility		0.173	0.131

4.5 Sex Selection Technology

Before we move to the role of public policies, it is illustrative to highlight the role of sex selection technology in the model. To this end, I assume that sex selection is prohibitively expensive. The outcomes of this experiment are shown in Table 10.

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 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

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 (% of household income)	5.70	5.70
Education expenditure per daughter (% of household income)	3.79	3.76
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

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 assess their importance relative to intrinsic sex preferences. After eliminating all these differences, the only remaining gender asymmetry is that mothers allocate their time between childcare, market, and home production, while fathers work full time in the market.

[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.

5.1 Economic Factors

To assess the effects of economic factors on the sex ratio, 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 their parents (given the number of sons and daughters) are determined by,

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

where each son or daughter has the same marginal impact on the level of transfers.
(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 (% of household income)	5.70	5.66	4.06	8.09	6.88
Education expenditure per daughter (% of household income)	3.79	4.49	2.42	5.87	5.58
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
Household income	100	98.8	85.7	125.1	114.7

Notes. The average household income is normalized to 100 for the benchmark economy.

The results show that not all economic factors have the same impact. They can increase or decrease the sex ratio and female education, depending on how they affect the trade-offs between sons and daughters and between the quantity and quality of children.

[34] This mimics the Dowry Prohibition Act 1961, which has not been enforced.

First, if the old-age support requirement is the same for sons and daughters, compared to the benchmark economy, daughters are more valuable for parents. As a result, 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%.

Second, without dowries, daughters become less costly. As a result, son-biased sex selections almost disappear, 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 cost of each birth is reduced and parents shift away from the quality to the quantity of children. As shown in the table, the fertility rate increases 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. As a result, the average household income decreases by 14.3%.

Third, if labor market discrimination is eliminated, the sex ratio at birth increases to 1.232. Why is this happening? One might expect that daughters become more valuable now as they have higher household income, but 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 now since their wives have higher incomes. 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, parents are more likely to end up with no sons 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 as a share of household income increases by 2.4 pp for sons and 2.1 pp for daughters, which leads the share of men and women

[35] Das Gupta et al. (2003) argue that increasing the 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.

with high education to increase by 22.0 pp and 26.6 pp, respectively. Due to increased educational attainment and female labor supply, the average household income increases by 25.1%.

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 more considerable extent by 22.9 pp, and the gender education gap declines significantly. Compared with the benchmark economy, the average household income is higher by 14.7%.

Overall, a combination of these factors can lead to an almost normal sex ratio and a small gender education gap, together with a large improvement in female labor supply, a substantial drop in fertility, and a large increase in household income. 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 set ω to 1, after making the economic factors symmetric for sons and daughters. The results are reported in Table 12.[37]

Once the gender differences in the economic factors are removed, further eliminating intrinsic sex preference has a limited effect on the SRB and the gender gap in education expenditure. The sex ratio is barely reduced, and the change in education expenditure as a share of household income is 0.8 pp for girls and -0.4 pp for boys, which is much smaller than the effects of the economic factors (1.8 pp and 1.2 pp). In addition, the effects on fertility, female labor supply, and household income 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

[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.

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 (% of household income)	5.70	6.88	6.46	5.33
Education expenditure per daughter (% of household income)	3.79	5.58	6.38	4.40
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
Household income	100	114.7	115.4	97.5

Notes. The average household income is normalized to 100 for the benchmark economy.

removing only the gender differences in the economic factors. In addition, its effects on fertility, female labor supply, and household income are rather small. The results suggest that the preference for sons can be largely attributed to gender differences in economic factors.

6 Can Policies Normalize the Sex Ratio?

How can governments correct 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. 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 for girls? How do they affect fertility, female labor supply, and household income?

In this subsection, I compare the impacts of these two policies. To this end, I introduce either a lump-sum subsidy for a female birth or a proportional subsidy for female education. For the birth subsidy, I assume that parents can receive a transfer for each girl. The value of the transfer is the same as in the *Dhanlakshmi* program, which is equivalent to 16.2% of the annual income of a male without secondary education. For the education subsidy, I assume that a fraction of education expenditure for daughters is paid by the government. The subsidy rate is chosen to obtain the same reduction in the SRB as caused by the subsidy for female births. The resulting subsidy rate is 18.8%. Both subsidies are financed by a proportional labor income tax. The outcome variables are reported in Table 13.[38]

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

	Benchmark	Birth subsidy	Education subsidy
Sex ratio at birth	1.139	1.094	1.094
Fertility	3.07	3.25	2.83
Education expenditure per son (% of household income)	5.70	5.33	5.59
Education expenditure per daughter (% of household income)	3.79	3.49	5.56
Men with high education (%)	64.6	60.6	65.1
Women with high education (%)	49.2	45.0	71.4
Gender gap in education	15.4	15.6	-6.3
Female labor (young)	0.250	0.235	0.303
Female labor (middle-aged)	0.270	0.277	0.313
Household income	100	95.2	102.1
Labor income tax rate (%)	0	0.45	0.91
Subsidies as a share of GDP (%)	0	0.42	0.84

Notes. The average household income is normalized to 100 for the benchmark economy.

Both policies reduce the SRB to 1.094 by design. The underlying mechanisms, however, are different. Subsidizing female births directly reduces the fixed cost of daughters,

[38] In Appendix D, I evaluate the impact of alternative subsidy amounts and show that the basic findings are largely unchanged.

while subsidizing female education makes educating daughters cheaper from the perspective of parents.

Despite having similar effects on sex selection, these two policies have different effects on female education because the parental quantity-quality trade-off is shifted in 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 60.6% and 45.0%, respectively. In contrast, the subsidy for female education directly motivates parents to increase their daughters' education expenditure, which increases the share of women with high education to 71.4%. Moreover, this leads to a lower fertility rate (2.83), 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 (65.1%).

In addition, 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 1.5 pp of their time, while subsidizing female education can increase it by 5.3 pp.

Since the two policies have opposite effects on educational attainment and female labor supply, they have different effects on household income. The subsidy for female births reduces the household income by 4.8%, while the subsidy for female education increases it by 2.1%.

Finally, the two policies imply different costs. A tax rate of 0.45% is required to finance the subsidy for female births, and the total subsidy expenditure is about 0.42% of GDP. For the subsidy for female education, the tax rate is 0.91%, and the expenditure is 0.84% of GDP, which is two times the expenditure on the subsidy for female births.

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, female labor supply, and household income. In contrast, a female education subsidy can lead to improvements in the SRB, fertility, female education, female labor supply, and household income. The monetary cost of the latter, however, is larger than that of the former.

6.2 Pensions

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, can be weakened if parents are covered by a pension system. In particular, the gender difference for old-age support would be less important. How much would a pension system affect parental sex preference? How could it affect other outcomes, including fertility, education investment, female labor supply, and household income?

To answer these questions, a pay-as-you-go pension system is introduced into the model. I assume that men receive a 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).[39] The results are shown in Table 14.[40]

A tax rate of about 14% is needed to finance the program. Since children become less valuable, parents now have slightly fewer children. 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 and women with high education is reduced to 56.6% and 44.4%, respectively. The average household income is reduced by 4.1%. 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.[41]

The analysis assumes that the pension system does not change the old-age support that the parents received from their children. Yet after a 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 change in social norms (Cheng et al., 2018; Mukherjee, 2020). To

[39] 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.

[40] See further results in Table A9 in Appendix D with different replacement rates.

[41] 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. Finally, 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.

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 (% of household income)	5.70	5.40	5.25	5.11
Education expenditure per daughter (% of household income)	3.79	3.84	3.91	4.05
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
Household income	100	95.9	95.3	95.0
Savings rate when middle-aged (%)	36.1	28.7	29.4	30.1
Labor income tax rate (%)	0	14.07	13.95	13.98

Notes. The average household income is normalized to 100 for the benchmark economy.

examine this possibility, I consider two hypothetical scenarios, 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 with high education increases by 0.5 pp, compared with the case with a 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

Parents in several South and East Asian countries want more sons than daughters and care more about their sons' education. The strong preference for sons results in prenatal sex screening and sex-selective abortions. As a result, the sex ratio at birth (the ratio of boys to girls) is well above the biologically normal ratio, and there are significant gender-education gaps. Yet, the relative importance of different factors behind son preference is not well understood, nor are their effects on fertility, female labor supply, and household income.

In this paper, I build an overlapping-generation model of fertility, sex selection, education investment, and marriage formation to quantify the causes and consequences of missing women in India. I focus on three economic factors that make daughter less valuable for parents: old age support provided by sons, the practice of dowry, and labor market discrimination against women.

The quantitative analysis reveals that economic factors play a much more critical role than the intrinsic sex preference of parents. In the benchmark model, the SRB is 1.14, and 65% of men complete secondary education, while only 49% of women achieve this. If there is no gender difference in the economic factors, the SRB drops to the biologically normal ratio. The total fertility rate declines from 3.1 to 2.4, and the share of women with secondary education rises significantly to 72%. Men also benefit from lower fertility, and the share of men with secondary education increases to 79%. In the benchmark economy, parents' incentives to have sons results in too many children, which lowers investment in boys and girls, but the impact on girls is much more substantial. Once these three economic factors are eliminated, eliminating intrinsic son preference has a minimal effect on the SRB.

The model is then used to evaluate policies that can change the sex ratio at birth. I compare subsidies for female births with subsidies for female education. Although both policies can reduce the sex ratio, they affect fertility, education, female labor supply, and household income differently by shifting the parental quantity-quality trade-off in opposite directions. A subsidy for each female birth of about 16% of the annual income of a male without secondary education reduces the SRB to 1.09. However, it raises the fertility rate to 3.3, reduces the share of men and women with secondary education to 61% and 45%, respectively, reduces young women's labor supply to 24% of their time, and reduces household income by 5%. A subsidy of 19% of education expenditure for girls could reduce the SRB to the same level. However, it reduces the fertility rate to 2.8, increases the share of women with secondary education substantially to 71%, increases young women's labor supply to 30% of their time, and increases household income by 2%.

Finally, introducing a pension system reduces sex selection and fertility. However, it also reduces children's education. A pay-as-you-go pension system with a 50% replacement rate reduces the SRB to 1.09. The shares of men and women with secondary education decline to 57% and 44%, respectively, since children's role as the providers of old-age support is now reduced. As a result, household income is reduced by 4.1%.

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., Choe, M. K., & Roy, T. K. (1998). Son preference, the family-building process and child mortality in India. *Population Studies*, *52*(3), 301–315.

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). 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. (1987). Selective discrimination against female children in rural Punjab, India. *Population and Development Review*, 77–100.

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.

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.

Hirway, I., & Jose, S. (2011). Understanding women's work using time-use statistics: The case of India. *Feminist Economics*, *17*(4), 67–92.

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

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.

Pongou, R. (2013). Why is infant mortality higher in boys than in girls? A new hypothesis based on preconception environment and evidence from a large sample of twins. *Demography*, *50*(2), 421–444.

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

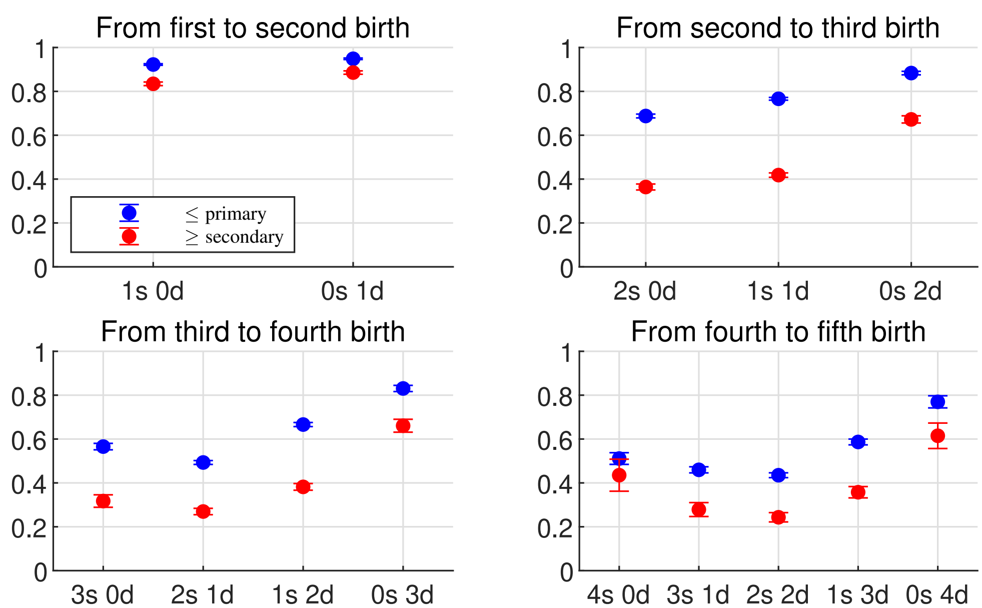
A Additional Empirical Facts

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.

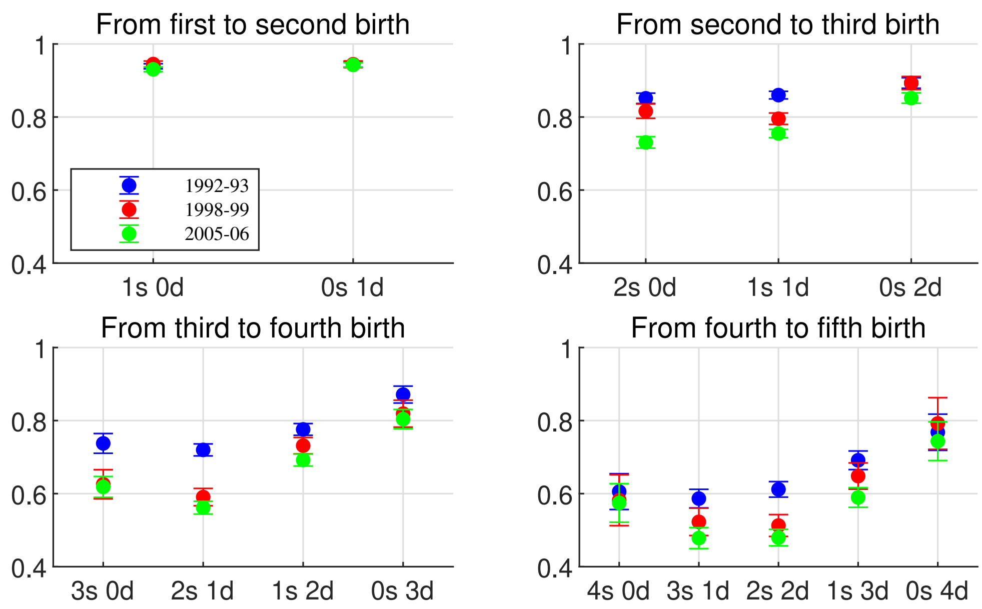
First, the analysis is repeated separately for women with and without secondary education. The results are plotted in Figure A1, suggesting that a son-biased fertility stopping rule is practiced by both groups.



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

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

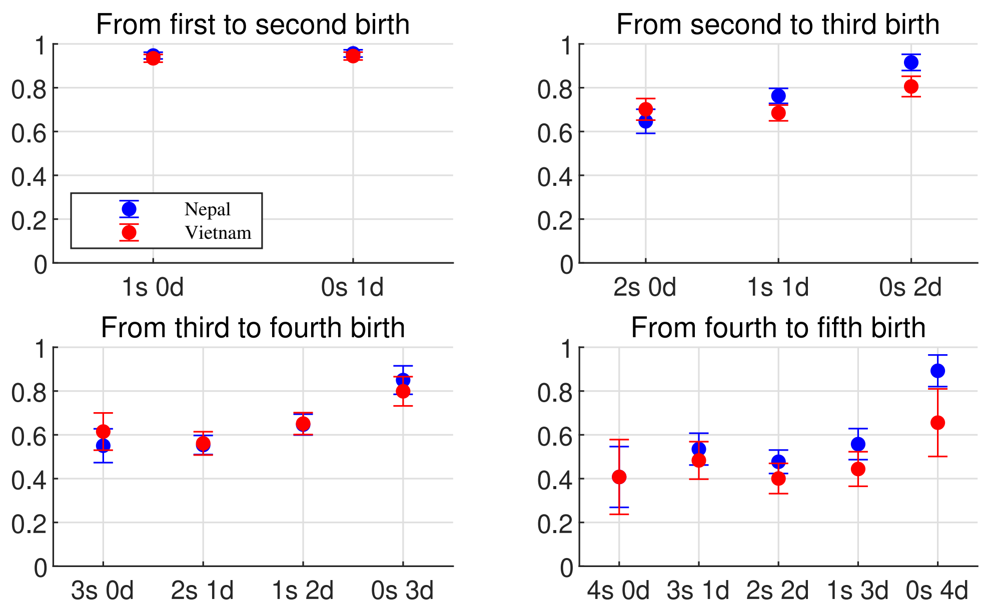
Second, the analysis is repeated for the previous waves of the survey in 1992-1993, 1998-1999, and 2005-2006. The results are plotted in Figure A2, suggesting that a son-biased fertility stopping rule has been persistent in India.



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 Existing 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 Existing Children in Nepal and Vietnam.

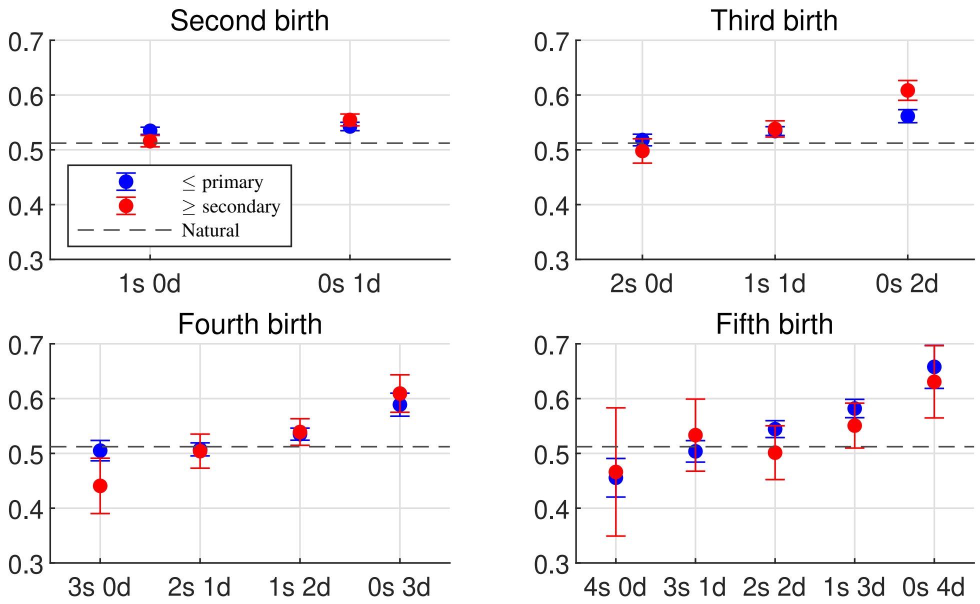
Fact 2. Son-biased Sex Selection.

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 levels. The results are plotted in Figure A4, suggesting that son-biased sex selections were practiced by both groups of women.

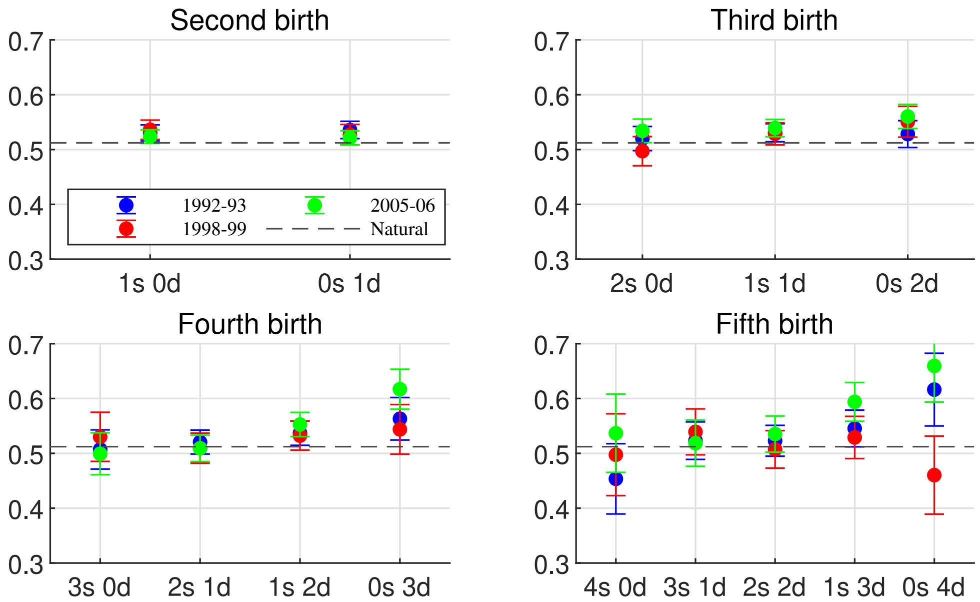
Second, the analysis is repeated for the previous waves of the survey in 1992-1993, 1998-1999, and 2005-2006. The results are shown in Figure A5, suggesting that sex selections were also practiced by women in the previous waves. One difference is that parents in these waves were less likely to select 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. However, 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.

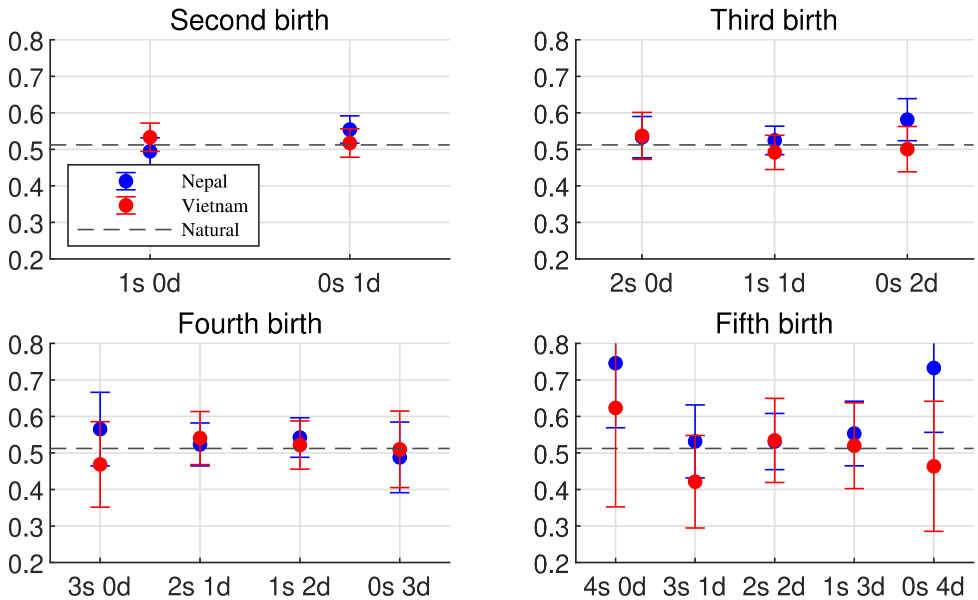
Figure A4. Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children by Women's Education



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

Figure A5. Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing 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 that the Next Birth is a Boy, Conditional on the Sex Composition of Existing 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 Model Estimation

B.1 Sensitivity Analysis of Estimation

This subsection 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 parameter.

B.2 Identification of the Substitution Elasticity between Sons and Daughters

This subsection shows the parity progression ratios and the sex selection decisions when the elasticity of substitution between sons and daughters, ϵ, is 1.

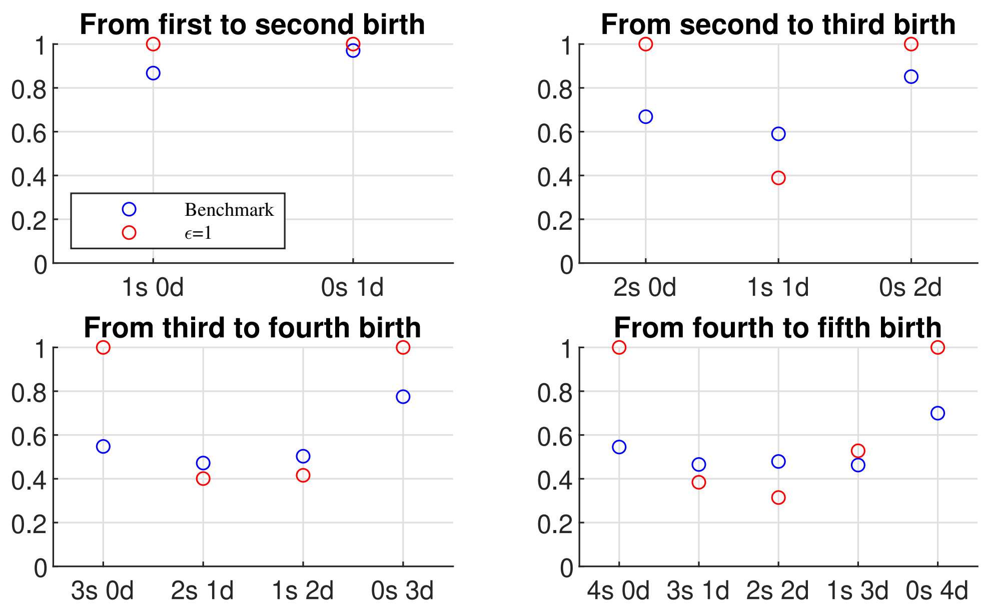


Figure A7. Parity Progression Ratio, Conditional on the Sex Composition of Existing Children (Benchmark and Unit Substitution Elasticity).

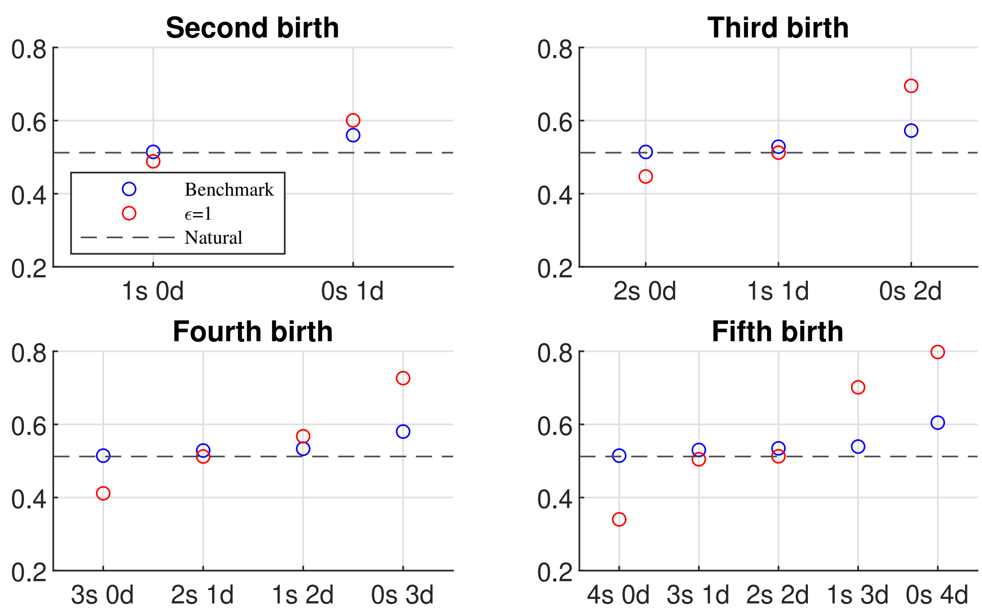


Figure A8. Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children (Benchmark and Unit Substitution Elasticity).

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 A9 and A10, respectively.

Table A6: Data and Model: Selected Moments for Punjab

Variable	Data	Model
Targeted moments		
Education expenditure per son (% of household income)	6.64	7.70
Education expenditure per daughter (% of household income)	5.75	5.40
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
Female labor supply when young	0.180	0.192
Female labor supply when middle-aged	0.167	0.159
Savings rate when middle-aged (%)	36.7	36.8
Aggregate moments		
Fertility	2.65	2.53
Sex ratio at birth	1.334	1.327

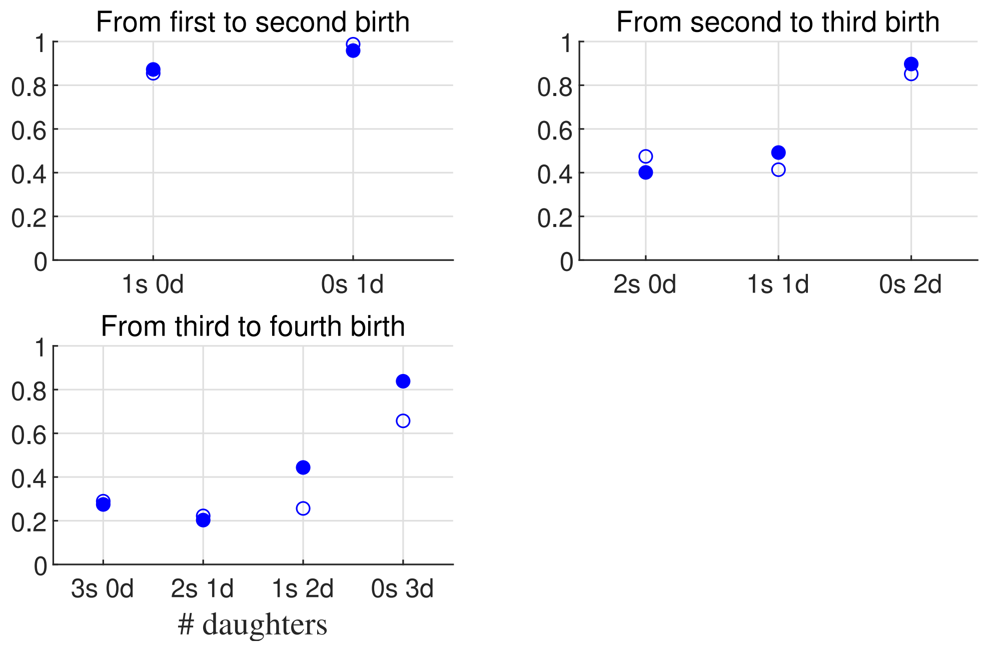


Figure A9. Parity Progression Ratio, Conditional on the Sex Composition of Previous Children in Punjab (Data and Model).

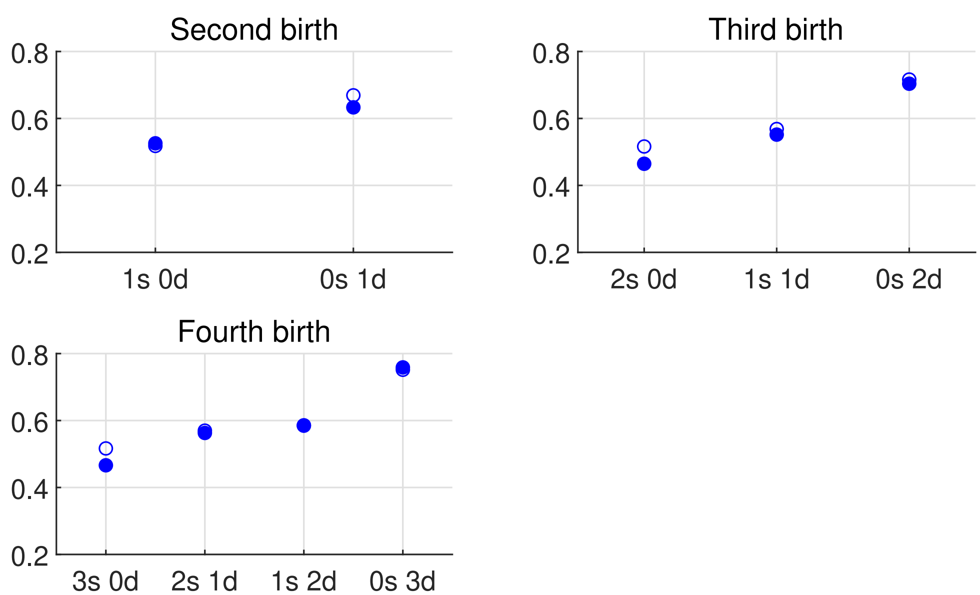


Figure A10. Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children in Punjab (Data and Model).

D Additional Results

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

Figure A11 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.

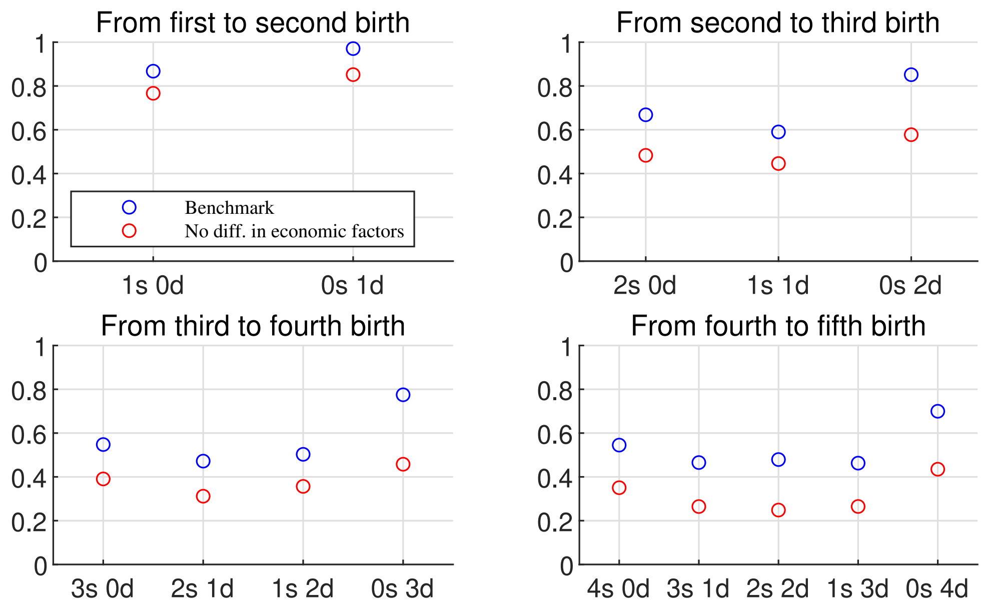


Figure A11. Parity Progression Ratio, Conditional on the Sex Composition of Previous Children if No Gender Difference in Economic Factors.

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 counterfactual economy almost do not practice son-biased sex selection even if they do not have sons.

Figure A13 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 A14 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.

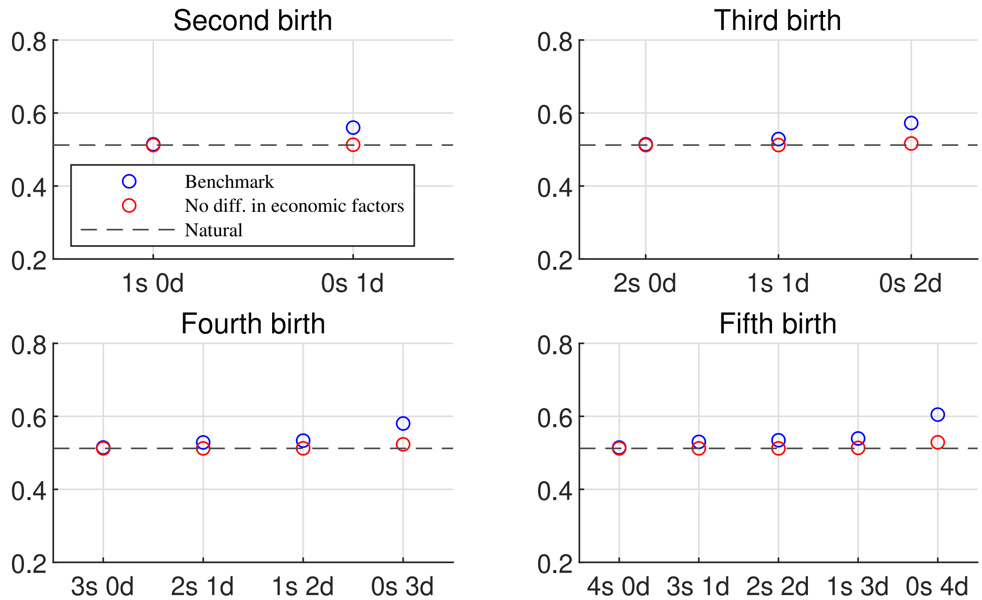


Figure A12. Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children if No Gender Difference in Economic Factors.

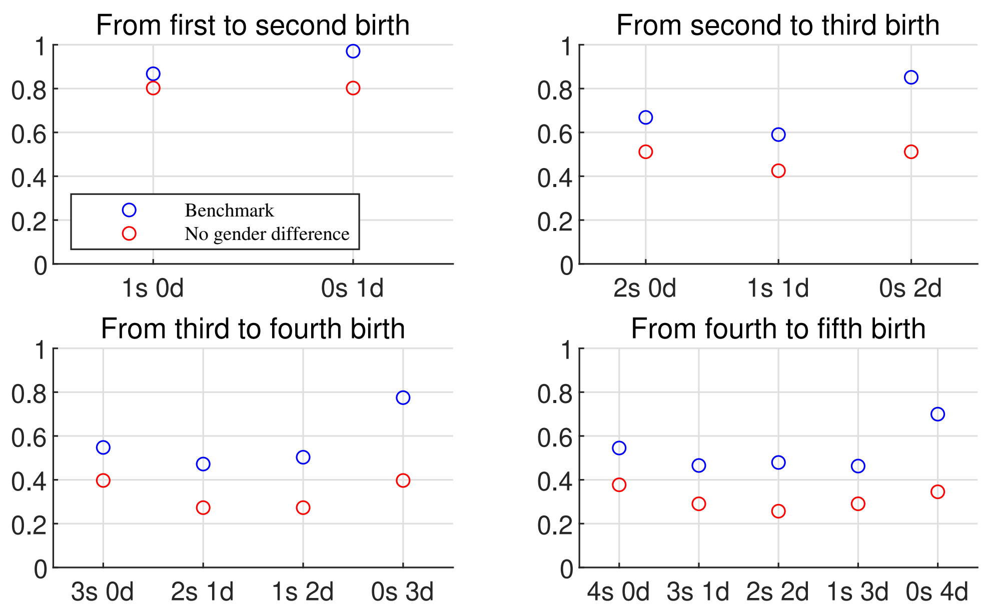


Figure A13. Parity Progression Ratio, Conditional on the Sex Composition of Existing Children if No Gender Difference.

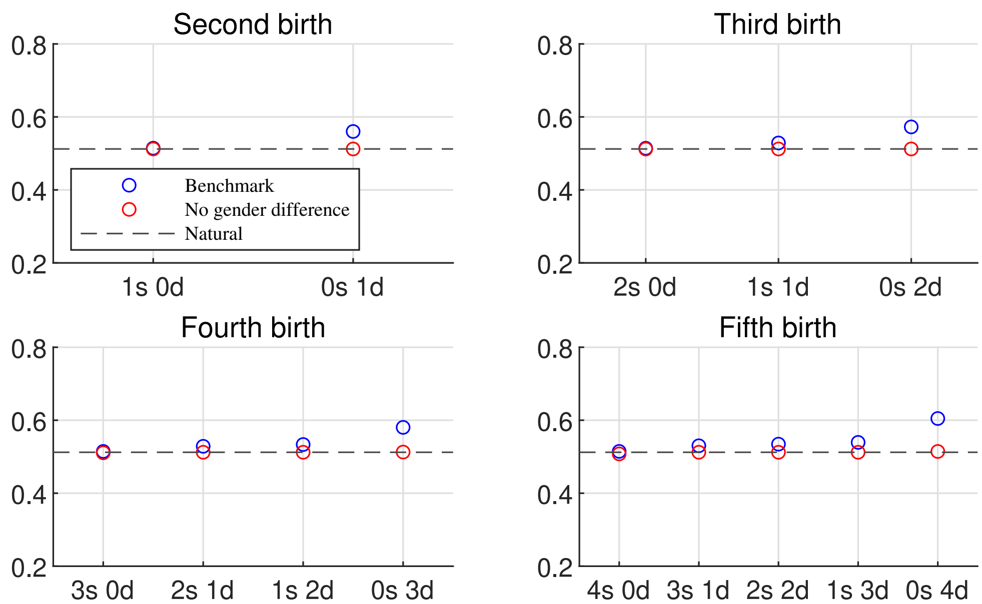


Figure A14. Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children if No Gender Difference.

D.2 Alternative 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 at 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 first 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	0.004	0.008	0.012
Sex ratio at birth	1.139	1.114	1.094	1.080
Fertility	3.07	3.14	3.24	3.33
Education expenditure per son (% of household income)	5.70	5.57	5.34	5.20
Education expenditure per daughter (% of household income)	3.79	3.67	3.50	3.35
Men with high education (%)	64.6	63.1	60.6	58.8
Women with high education (%)	49.2	47.6	45.2	43.0
Gender gap in education	15.4	15.5	15.5	15.8
Female labor (young)	0.250	0.245	0.235	0.227
Female labor (middle-aged)	0.270	0.273	0.277	0.279
Household income	100	97.7	95.3	93.4
Labor income tax rate (%)	0	0.20	0.44	0.71
Subsidies as a share of GDP (%)	0	0.14	0.41	0.66

Notes. The subsidies are equivalent to 8%, 16%, and 24% of the annual income of a male without secondary education, respectively. The average household income is normalized to 100 for the benchmark economy.

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 (% of household income)	5.70	5.72	5.56	5.24
Education expenditure per daughter (% of household income)	3.79	4.72	5.68	6.69
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
Household income	100	101.2	102.2	102.7
Labor income tax rate (%)	0	0.38	1.00	1.97
Subsidies as a share of GDP (%)	0	0.35	0.92	1.81

Notes. The average household income is normalized to 100 for the benchmark economy.

D.3 Alternative Replacement Rates of Pensions

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 (% of household income)	5.70	5.42	5.40	5.38
Education expenditure per daughter (% of household income)	3.79	3.89	3.84	3.81
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
Household income	100	97.0	95.9	95.0
Savings rate when middle-aged (%)	36.1	32.4	28.7	24.5
Labor income tax rate (%)	0	8.46	14.07	19.79

Notes. The average household income is normalized to 100 for the benchmark economy.