# Missing Women: A Quantitative Analysis*

Yongkun Yin

Shandong University

The sex ratio at birth (SRB) is abnormally high in China, India, and some other developing countries. Why do parents favor sons over daughters? What policies can normalize the sex ratio? To answer these questions, I develop an overlapping-generation model of fertility, sex selection, and the quantity-quality trade-off and estimate it for India. The analysis shows that the main drivers of missing women are economic factors: old-age support from sons, dowry for daughters, and gender discrimination in the labor market. Equalizing these factors across genders would reduce the SRB from 1.12 to 1.05, reduce fertility from 3.1 to 2.2, and significantly improve children's education. Subsidies for female births and female education can 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. A pay-as-you-go pension system can reduce the SRB, but it also reduces children's education.

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

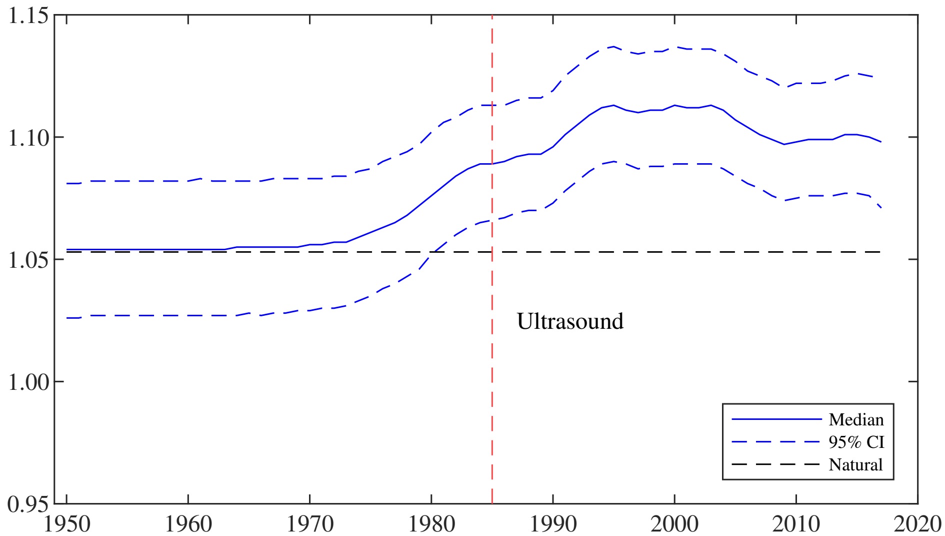
*JEL codes*: D1, J13, J16.

---

*Yongkun Yin, School of Economics Shandong University, Jinan 250100, China. Email: yongkunyin@gmail.com. I am grateful to Nezih Guner for his support and guidance at different stages of the project. I thank Sonia Bhalotra, Samuel Bentolila, Tiago Cavalcanti, Sebastián Fanelli, Raquel Fernández, Luisa Fuster, Paula Gobbi, Pedro Mira, Josep Pijoan-Mas, Alessandro Ruggieri, Tom Zohar, and seminar participants at CEMFI, Shaanxi Normal University, Jinan University, Liaoning University, Hunan University, Shanghai University of Finance and Economics, the 2023 Shandong University Summer Workshop in Economics, the 24th Labor Economist Quarterly Forum, the 2024 Shandong University Spring Workshop in Applied Microeconomics, the 2024 AMES (China), the 2024 SEHO, and the World Bank for insightful comments and discussions. I thank Jipeng Zhang for providing high-performance computation. I acknowledge the funding from Fundación Ramón Areces.

# 1 Introduction

In several South and East Asian countries, the male-to-female ratio is disproportionately high, a phenomenon referred to as "missing women" by Sen (1990). Figure 1 shows the sex ratio at birth (SRB) in India since the 1950s. The natural SRB is around 1.05, which remains stable across time and space for biological reasons. However, India's SRB began to rise sharply in the 1980s, coinciding with the increased availability of ultrasound technology that enabled prenatal sex detection. By the 1990s, the SRB had exceeded 110 male births per 100 female births, significantly above the natural level.




*Notes.* Data are from Chao et al., 2019.

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

A skewed SRB is not unique to India. Across various developing regions, parents want more sons than daughters and aim to have at least one son. In China, India, and some other Asian countries, the preferences for sons have led to sex-selective abortions, higher female child mortality, and lower female educational attainment (Chao et al., 2019; Barcellos et al., 2014).

Why are there missing women? Most studies attribute the preferences for sons to the gendered benefits and costs of having children, such as old-age support predominantly provided by sons (Das Gupta et al., 2003; Chung and Das Gupta, 2007) and dowry payment for daughters (Bhalotra et al., 2020). Daughters may also be devalued due to their disadvantages in the labor market (Qian, 2008; Carranza, 2014). In addition, intrinsic son preference, where parents derive more inherent utility from having sons, may also play a role.[1] Despite these insights, there is limited understanding of the relative importance of these factors, their interactions, and the most effective policies to address the skewed sex ratio.

This study addresses the following questions: What are the relative contributions of different factors to the imbalanced sex ratio? In particular, how do economic factors—old-age support, dowry payment, and gender discrimination in the labor market—compare to intrinsic son preferences? How do these factors influence fertility rates, children's education, women's labor supply, and household income? Moreover, what policies can normalize the sex ratio, and how would they affect related outcomes?

To answer these questions, I develop a quantitative overlapping-generation (OLG) model that incorporates various motives for prenatal sex selection, within a framework that includes parental quantity-quality trade-off and marriages. In this model, parents pay dowries for their married daughters and rely on their adult children, mainly sons, for old-age support. There is gender discrimination in the labor market, which creates a wedge between women's productivity and wages. These economic factors influence parental desires for the quantity and quality of sons versus daughters. Additionally, parents may have an intrinsic preference for sons. Fertility decisions are made sequentially, conditional on the sex composition of existing children. Parents may use sex selection technology to ensure a desired sex at birth, incurring both monetary and utility costs. Parental decisions on fertility, sex selection, and educational investment are forward-looking, taking into account how these decisions affect children's marriage prospects, dowry payment, and old-age support received from children.

The model parameters are calibrated using Indian data. The model can replicate

[1] Intrinsic son preference may result from evolutionary pressures (Trivers and Willard, 1973), patrilineal cultures emphasizing lineage continuity and ancestor worship (Das Gupta et al., 2003), or prolonged exposure to such cultural norms (Almond et al., 2013).

key moments conventionally used to discipline OLG models of fertility and investment in children, such as women's working hours, the education gradient in fertility, and education expenses per son and daughter. Notably, the model can also replicate observed probabilities of having another child and the likelihood that this child is male, conditional on the sex composition of existing children. These moments are critical for identifying the substitutability of sons and daughters in parental preferences and the strength of intrinsic son preferences.

Using the estimated parameters, I first conduct validation exercises by simulating household responses to changes in dowry values and cash transfer programs for families with daughters. The model's predicted elasticity of sex ratios with respect to dowry values is consistent with the estimate of Bhalotra et al. (2020), who use exogenous gold price variations to study the impact of dowry on sex selection. The model also predicts that the *Dhanlakshmi* program, which subsidized female births and education in seven Indian states from 2008 to 2013, would increase the likelihood of female births, with the magnitude closely matching the difference-in-differences estimate of Biswas et al. (2023).

I then conduct counterfactual experiments to quantify the contributions of different factors to the imbalanced sex ratio and their effects on fertility, education, female labor supply, and household income. I first remove gender differences in one economic factor at a time, and then all at once, by (a) requiring daughters to provide the same old-age support as sons, (b) eliminating gender discrimination in the labor market, and (c) reducing dowry payments by half and assuming equal marriage payments for sons. Finally, I eliminate intrinsic sex preferences.

The results reveal that economic factors significantly affect the SRB, more so than intrinsic son preferences. In the benchmark model, the SRB is 1.122. Removing gender differences in economic factors reduces it to the normal level of 1.050. Eliminating intrinsic sex preferences has little further effect.

Interestingly, gender discrimination in the labor market alone does not drive sex selection. Indeed, removing this factor increases the SRB to 1.220, rather than reducing it. This is because the value of daughters does not increase unless their income contributes to old-age support, while the value of sons increases due to their wives' higher earnings (Das Gupta et al., 2003). Moreover, as women work more and have fewer children, the likelihood of not having sons increases without sex selection, prompting parents to select male births (Jayachandran, 2017).

Economic factors also have large effects on fertility, education, female labor supply, and household income. In the benchmark model, the fertility rate is 3.1, with 70.7% of men and 53.0% of women completing secondary education. Women devote only 24.5% of their time to market work. Removing gender differences in economic factors reduces the fertility rate to 2.2, increases the secondary education completion rate to 90.0% for men and 84.6% for women, and increases women's market work to 45.2% of their time. These changes lead to a 27.2% increase in average household income.

Next, I evaluate two policies aimed at normalizing the sex ratio: a subsidy for female births and a subsidy for female education, both commonly used by Indian state governments. Although both policies can reduce the sex ratio, they have different effects on female education, fertility, and female labor supply due to their distinct impacts on parental quantity-quality trade-offs. A female birth subsidy increases fertility, reduces education per child, and decreases female labor supply, while a female education subsidy has the opposite effects.

Finally, I assess the impact of a pension system. As an alternative source of post-retirement income, pensions reduce the reliance on children for old-age support and thus lower the SRB. A pay-as-you-go pension system with a replacement rate of 50%, the target rate of the Employees Provident Fund Scheme (the largest pension scheme in India), would reduce the SRB to 1.095, but it would also reduce the secondary education completion rate to 65.3% for men and 46.2% for women.

This paper builds on the macro-development literature on fertility and child investment, grounded in the quantity-quality trade-off theory (Becker and Lewis, 1973; Becker, 1981). While previous studies have explored the effects of technological progress (Galor and Weil, 2000; Delventhal et al., 2021), child mortality (Kalemli-Ozcan, 2002, 2003), and adult mortality (Soares, 2005; de la Croix and Licandro, 2013), my focus is on the drivers of sex selection. Furthermore, recent studies by Cavalcanti et al. (2021)

and Seshadri and Zhou (2022) have examined how family planning policies that subsidize the use of contraceptive technology can reduce fertility and boost education. In contrast, I examine how gender-equalizing policies aimed at counteracting prenatal sex selection technology affect sex ratios, fertility, and human capital.

This study also intersects with economic models on the demographic consequences of sex selection technology. Leung (1994) and Davies and Zhang (1997) assume static, one-time fertility choices, while Li and Pantano (2023) assume dynamic, sequential fertility choices, as my model does. Unlike Li and Pantano (2023), my model considers a broader range of motives for sex selection beyond intrinsic preferences, incorporates education and female labor supply decisions to capture their interactions with fertility decisions, and allows for feedback from children's marriages into parental sex selection decisions.

This paper contributes to the economic and demographic literature on the causes of skewed SRB. Existing studies often focus on specific aspects, such as old-age support provided by sons (e.g., Das Gupta et al., 2003), dowry payments for daughters (e.g., Bhalotra et al., 2020), labor market disadvantages faced by women (Qian, 2008; Carranza, 2014), and intrinsic sex preferences (e.g., Abrevaya, 2009; Almond et al., 2013), while my research employs a quantitative model to assess the relative importance of multiple factors and examine their interactions.

This paper also contributes to the literature evaluating policies targeting sex ratios. In particular, Nandi and Deolalikar (2013) and Rastogi and Sharma (2022) examine the impact of banning sex selection on the SRB, while Anukriti (2018) and Biswas et al. (2023) examine how cash transfer programs for families with daughters affect the SRB and fertility. These studies typically use empirical methods to estimate the partial equilibrium and short-run effects of small-scale policies. In contrast, my quantitative OLG model can be used to assess the long-run effects at the aggregate level.

The remainder of the paper is structured as follows. Section 2 presents the empirical motivations for the study. Section 3 describes the model, followed by its calibration in Section 4. Section 5 compares the model's predictions with quasi-experimental evidence. Section 6 investigates the causes and consequences of missing women. Section

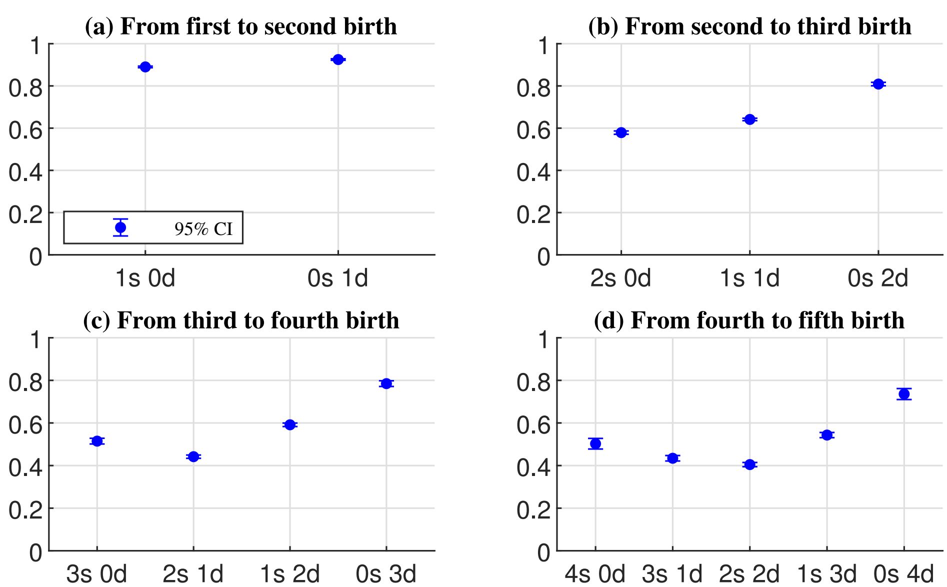
7 evaluates policies to address the imbalanced sex ratio. Section 8 concludes.

# 2 Facts

In this section, I demonstrate how the decision to have another child and control the sex of this child is influenced by the sex composition of existing children, using data from the Indian Demographic and Health Survey (DHS). The Indian DHS was a repeated cross-sectional survey conducted in 1992-1993, 1997-1998, 2005-2006, and 2015-2016. Women aged 15-49 years reported their complete fertility history, detailing the sex and date of each birth. I focus on women aged 40-49 years, as they are likely to have completed their fertility, and restrict the sample to those with 1-6 children, as women with fewer or more children are rare. The main analysis uses the 2015-2016 data, with earlier waves used for robustness checks.

***Fact 1. Son-biased Fertility Stopping Rule***

I first calculate the parity progression ratio (PPR), or the proportion of women who progress from one parity to the next, conditional on the sex composition of their current children. The ratios are displayed in Figure 2.




*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.

An important pattern is that *parents are more likely to have another child if they have fewer sons than daughters, demonstrating a son-biased fertility stopping rule.* For example, among women with two children, 80.9% will have another child if they have two daughters and no sons, compared to 57.9% if they have two sons and no daughters (see Panel (b)). Among women with three children, 78.5% will have another child if they have three daughters and no sons, compared to 51.5% if they have three sons and no daughters (see Panel (c)). This asymmetry in PPR clearly indicates a desire for sons.

In addition, parents are more likely to have another child if the sex composition of their existing children is imbalanced, either with too few sons or too few daughters. This pattern is particularly evident among women with three or four children.[2] For example, among women with four children, only 40.5% will have another child if they have two sons and two daughters. However, this proportion increases to 50.3% if they have only sons and to 73.6% if they have only daughters (see Panel (d)). This U-shaped PPR suggests that parents do not view sons and daughters as perfect substitutes and value a balanced sex composition.

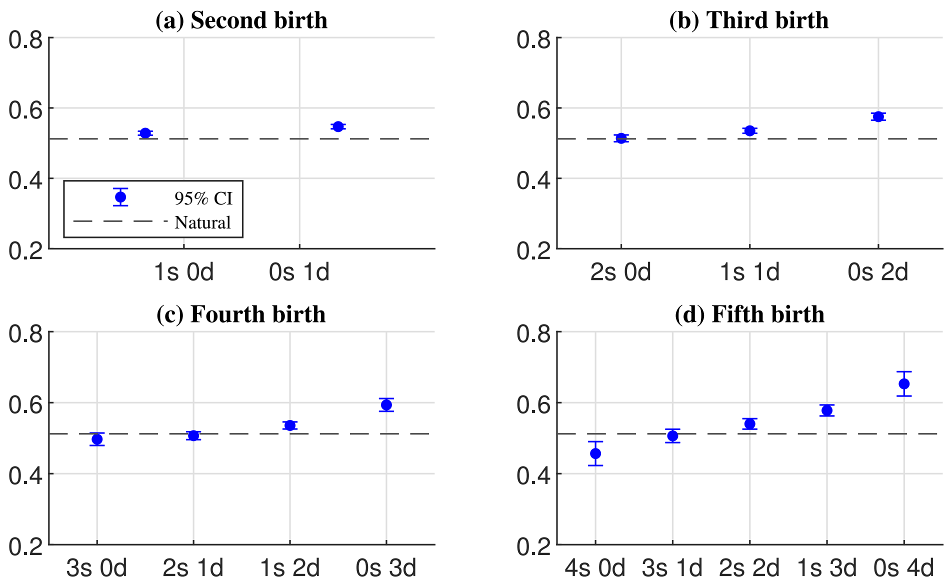
Extensive robustness checks are conducted for these findings. The patterns hold across women with different levels of education and across castes. They are also consistent when analyzed using earlier data in 1992-1993, 1998-1999, and 2005-2006. Additionally, similar trends are observed in Nepal and Vietnam, where many parents express a preference for sons over daughters. Details are shown in Appendix A.

***Fact 2. Son-biased Sex Selection***

I now examine how the sex composition of existing children influences the decision to select the sex of the next child. Using data from the 2015-2016 Indian DHS, I calculate the probability that the next birth will be male, conditional on the sex composition of existing children. The results are shown in Figure 3.

A striking pattern is that *the probability of the next birth being male increases as*

[2]This pattern is not observed for women with two children, as the progression ratio is lower for those with two daughters compared to those with a son and a daughter. In this case, the desire for sons outweighs the desire for a balanced sex composition.




*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.

*parents have fewer sons, a trend that is consistent across all birth orders.* When sons make up a large proportion of existing children, this probability remains close to the natural rate of 51.2%. However, when sons are few, it increases significantly. For example, among women having a third child, 51.4% of the third births are male if the first two children are boys. However, this probability increases to 53.5% if the first two are a boy and a girl, and to 57.5% if the first two are both girls (see Panel (b)). The deviations from the natural probability suggest sex-selective abortions favoring male births.

Parents may also select a female child to achieve a balanced sex composition. For example, among women who have a fifth child after four sons, only 45.7% of the fifth births are boys, 5.5 percentage points below the natural probability (see Panel (d)). Despite these cases, son-biased sex selection remains the dominant practice.

Extensive robustness checks are conducted for these findings. The patterns hold across women with different levels of education and across castes. They are also consistent when analyzed using earlier survey data in 1992-1993, 1998-1999, and 2005-2006. However, sex selection was less common in earlier survey waves due to limited access to sex-selection technology for older cohorts. Additionally, while sex selection is evident in the 2016 Nepalese DHS data, it is not detectable in the 2002 Vietnamese DHS data, likely due to technological constraints in Vietnam during the corresponding period. Details are shown in Appendix A.

To sum up, Facts 1 and 2 reveal that Indian parents are more likely to have another child, and this child is more likely to be male, when they have fewer sons. This suggests a clear desire for more sons than daughters and, in many cases, a desire for at least one son. As will become clear, these patterns are crucial for identifying the strength of intrinsic son preferences in the model.

***Other Facts on the Quantity and Quality of Children***

In Appendix A, I present additional well-known facts about the quantity and quality of children. First, there is a negative correlation between the number of children and their educational attainment. Second, there is a significant gender gap in education, with girls receiving on average two fewer years of education than boys given the same family composition. Third, parental education is inversely related to family size, with maternal education having a particularly strong impact.

# 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 live through four periods: period 0 as children, period 1 as young adults, period 2 as middle-aged adults, and period 3 as elderly adults. Each period corresponds to 20 years. There are two sexes in the model, male and female, denoted by $g \in \{f, m\}$.

Children do not make any economic decisions. They receive education and acquire human capital. Individuals can attain one of two levels of education/human capital: low or high, denoted by $h_g \in \{h^L, h^H\}$, which is determined by parental investment in education. When children become young adults, they get married. Newlywed couples receive a dowry from the bride's parents. Young married couples have children, invest in their children's education, work, and consume. Parents can select the sex of their children at a monetary and utility cost.

Middle-aged adults work, consume, and save for their old age. They also pay dowries to their young married daughters, if they have any. Additionally, middle-aged households provide financial transfers to their elderly parents. These transfers depend on the number of siblings of the husbands and wives and their sex composition. Elderly adults do not work. They consume their savings and transfers received from their children. The life cycle is illustrated in Figure 4. Next, I describe the model in detail.

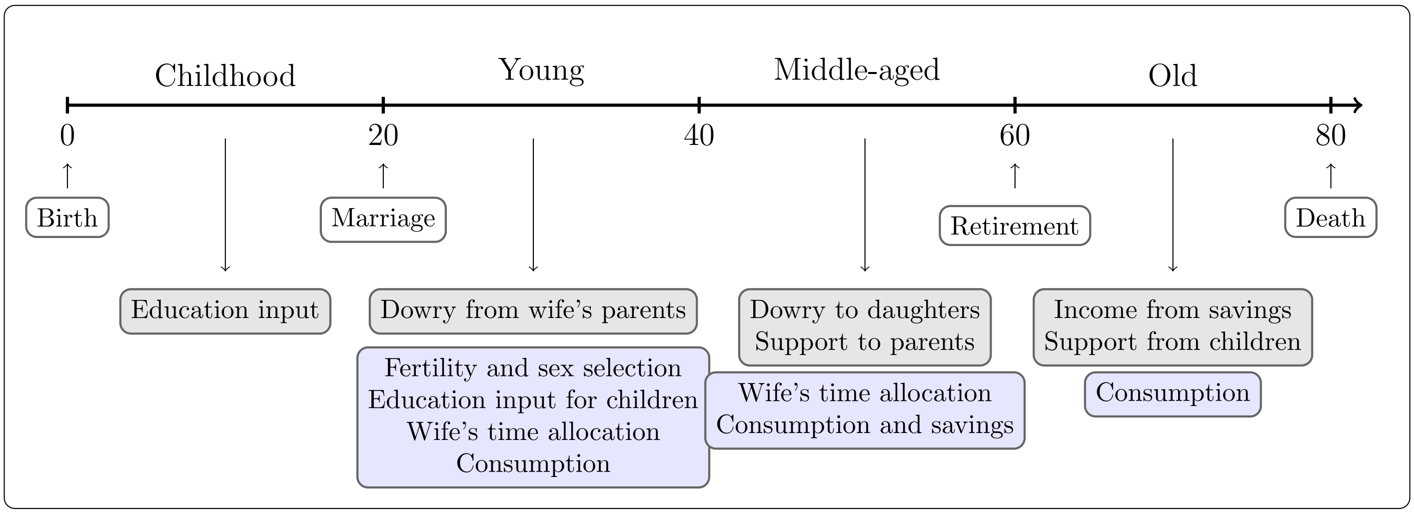



Figure 4: Life Cycle

## 3.1 Households

***Young Married Households***

Consider a young couple with human capital levels $\{h_m, h_f\}$. Each spouse is endowed with one unit of time. The husband works exclusively in the labor market, while the wife divides her time between market work ($l_1$), household production, and childcare, in line with time-use data (Hirway and Jose, 2011). Following Cavalcanti and Tavares (2016) and Hsieh et al. (2019), women face discrimination in the labor market, reflected in a gap between their wage rates and productivity. I assume that the cost of this discrimination is redistributed as a lump sum transfer, $T$, to working households. Upon marriage, the young couple receives a dowry ($d$) from the wife's parents. The household income is given by,

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

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

Starting with no children, the young couple decides whether to have a child, $b \in \{0,1\}$. For each birth, the parents also decide whether to use technology to select a male or female child ($s_m, s_f \in \{0,1\}$). Without sex selection, each child can be male with probability $p_m$ and female with probability $1 - p_m$. With sex selection, parents can ensure the child's sex. However, they must pay a monetary cost $\pi$ and incur a utility cost $\xi$. The monetary cost represents the price of prenatal sex detection and sex-selective abortion, while the utility cost reflects the psychological and physical side effects of abortion, which vary among individuals. I assume that $\xi$ is randomly drawn from a log-normal distribution after marriage, i.e., $\log(\xi) \sim N(\mu_\xi, \sigma_\xi)$.

After the first child is born and the sex is known, parents decide whether to have another child and select the sex of this child. Thus, fertility and sex selection decisions are made sequentially, conditional on the sex composition of existing children.

Suppose that the parents end up with $q_m$ male children and $q_f$ female children, among which $q_s$ were born by sex selection. Parents then decide how much to spend on education for each son and daughter ($e_m$ and $e_f$). 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 requires a time commitment from the mother, denoted by $\tau$. The time not spent on market work and raising children is used for domestic production using a linear 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, deriving utility from both the number of children and their educational investment. However, they differentiate between sons and daughters. The utility derived from children is given by,

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

The parameter $\alpha$ represents the level of altruism, randomly drawn from a log-normal distribution after 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 children's basic human capital and is independent of parental investment. The parameter $\eta$ governs the relative importance of the quality (education investment) versus the quantity of children. The parameter $\omega$ reflects the relative importance of sons versus daughters: $\omega \in (0.5, 1]$ indicates an intrinsic preference for sons, while $\omega \in [0, 0.5)$ indicates a preference for daughters. The quantity and quality of sons and daughters are aggregated using a constant-elasticity-of-substitution (CES) function, with $\epsilon$ governing the elasticity of substitution. If $\epsilon > 1$, sons and daughters are substitutes, and if $\epsilon < 1$, they are complements. Setting $\omega = 0.5$, $\epsilon = +\infty$, and $e_m = e_f$ yields a standard functional form without sex differentiation, which is commonly used in the literature (e.g., de la Croix and Doepke, 2003; Cavalcanti et al., 2021; Delventhal et al., 2021).

Whether children can achieve a high level of education is uncertain, and its probability depends on the education spending,

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

where $z > 0$ governs the efficiency of human capital production. This probability increases with the education spending, approaching zero as the spending goes to zero, and approaching one as it goes to $+\infty$.

***Middle-Aged Married Households***

Each spouse is endowed with one unit of time. The husband works exclusively in the labor market, while the wife divides her time between market work ($l_2$) and household production ($n_2 = 1 - l_2$). In addition, they receive a lump-sum transfer, $T$, which is the redistributed discrimination cost. The couple decides on their savings ($a$) for old age.[3]

If the couple has daughters, each marries during this period with a dowry equal to an exogenous fraction ($\delta$) of the father's income.[4,5] Therefore, household income net

[3]The assumption that only middle-aged parents can save is made for computational tractability.

[4]The determinants of dowry values remain unclear (Anderson, 2007). While dowries are often viewed as a price to clear the marriage market (Becker, 1981), empirical studies show varying effects of the sex ratio (Rao, 1993; Edlund, 2000) and the groom and bride's education (Behrman et al., 1999; Maertens and Chari, 2020; Calvi and Keskar, 2021; Chiplunkar and Weaver, 2023). However, a robust finding is that dowries are positively associated with parental income or wealth (Anderson, 2007; Maertens and Chari, 2020).

[5]Assuming that the dowry payment is proportional to the parents' income, rather than the father's income alone, will make the model less tractable.

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 make monetary transfers to their elderly parents. The transfer to the husband's parents is an exogenous fraction, $t_m$, of their income net of dowry and savings ($I_2$), while for the wife's parents, it is $t_f$. As detailed later, $t_m$ 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 couple's consumption of market goods in middle age 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***

Elderly parents do not work and rely on 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 (net of dowries and savings), while a daughter transfers $t_f$. Both $t_m$ and $t_f$ depend on the number of sons and daughters, as given by

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

and,

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

The parameter $\theta_m$ is the transfer level when parents have only one son, and $\kappa_m \in (0,1)$ governs how transfers per son change with the number of sons. As the number of sons increases, transfers per son decrease, but total transfers can rise. Similarly, $\theta_f$ is the transfer level when parents have only one daughter, and $\kappa_f \in (0,1)$ governs how the transfers per daughter change with the number of daughters.

A key feature of this specification is that daughters do not provide old-age support unless they have no brothers, capturing the fact that that elderly parents typically receive support from their sons through coresidence.[6] The interpretation here is that

[6]Data from the Longitudinal Ageing Study in India in 2017-2018 suggest that less than 5% of

when parents and children coreside, a fraction of the children's income can be effectively transferred to the parents. In the calibration section, I will map monetary transfers to coresidence and demonstrate how this specification replicates precisely the pattern of coresidence.

The consumption in old age is given by,

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

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

***Single Households***

In the model, there can be some single-male households. Consider a single male with human capital $h_m$. When young and middle-aged, he earns income $h_m$ and receives a lump-sum transfer of $T$. He does not engage in home production. When young, his income is fully consumed. In middle age, he saves $a$ for old age, with a fraction of the remaining income transferred to his parents, as specified by Equation (8). In old age, he consumes his savings, $a(1+r)$.

## 3.2 Problems of Married Households

The household problems are solved using backward induction. I focus on married households, as the problems of single households are straightforward. I start by computing the value of being elderly and then examine the decisions regarding savings and female labor supply of middle-aged households. Next, I solve for the decisions on education and female labor supply of young households, given the number of sons, daughters, and sex-selected children. Finally, I examine the sequential fertility and sex selection decisions.

***Savings, Female Labor Supply, and Education Investment Decisions***

Elderly households make no economic decisions. They consume their savings and transfers from their children, which is determined by the current state vector $\mathbf{x}_3 = \{a, q_m, q_f, \mathbf{x}_2^{k,m}, \mathbf{x}_2^{k,f}\}$. Here, $\mathbf{x}_2^{k,m}$ and $\mathbf{x}_2^{k,f}$ represent the state of their male and female

individuals aged above 60 years received monetary transfers from their children, but more than 60% were residing with them.

children in middle age. Children's states affect parents' decisions due to intergenerational linkages. Therefore, the value of being elderly 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 middle-aged households, the state vector is $\mathbf{x}_2 = \{h_m, h_f, q_m, q_f, \mathbf{x}_1^{k,m}, \mathbf{x}_1^{k,f}, \mathbf{x}_3^{p,m}, \mathbf{x}_3^{p,f}\}$. Here, $\mathbf{x}_1^{k,m}$ and $\mathbf{x}_1^{k,f}$ represent the state of their young male and female children, respectively, and $\mathbf{x}_3^{p,m}$ and $\mathbf{x}_3^{p,f}$ represent the state of the husband's and wife's parents, respectively. Clearly, $\mathbf{x}_3^{p,m}$ includes $\{q_m^{p,m}, q_f^{p,m}\}$, i.e., the number of male and female children in the natal family of the husband, while $\mathbf{x}_3^{p,f}$ includes $\{q_m^{p,f}, q_f^{p,f}\}$, i.e., the number of male and female children in the natal family of the wife.

Given $\mathbf{x}_2$, middle-aged households choose $\{l_2, n_2, c_2, a\}$ to solve the following problem,

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

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, the parameter $\nu_2$ governs the relative importance of non-market versus market goods consumption in middle age, and $\beta$ is the discount factor. Notice that when middle-aged households predict their future state ($\mathbf{x}_3$), they must predict the future state of their children ($\{\mathbf{x}_2^{k,m}, \mathbf{x}_2^{k,f}\}$), as this is part of $\mathbf{x}_3$.

Now, consider the problem of young households. First, suppose that the number of sons, daughters, and sex-selected children, $\{q_m, q_f, q_s\}$, has been determined. Then, parents only decide on education investment and female labor supply. The state vector is $\mathbf{x}_1 = \{h_m, h_f, q_m, q_f, q_s, \mathbf{x}_2^{p,f}, \alpha, \xi\}$, where $\mathbf{x}_2^{p,f}$ is the current state of the wife's parents, which determines dowries received by the young couple. Given $\mathbf{x}_1$, young households choose $\{l_1, n_1, e_m, e_f, c_1\}$ 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

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

and

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

In the objective function, the parameter $\nu_1$ governs the relative importance of non-market versus market goods consumption when young. The value of $\nu_1$ can differ from that of $\nu_2$, as people may have varying tastes for domestic goods at different ages. Note that, in the budget constraint, $d(\mathbf{x}_2^{p,f})$ is the dowry payments from the wife's parents.

Now we have obtained the value of having $q_m$ sons and $q_f$ daughters, with $q_s$ children born using sex selection technology. Note that the other variables in the state vector $\mathbf{x}_1$, except $\{q_m, q_f, q_s\}$, are beyond the household's control. For clarity, in the following, I will use $V_1(q_m, q_f, q_s)$ instead of $V_1(\mathbf{x}_1)$ to denote the value of ending up with $\{q_m, q_f, q_s\}$.

***Fertility and Sex Selection Decisions***

Now consider parental decisions on fertility and sex selection, which are made in a sequential manner, 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 is 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 simply $V_1(q_m, q_f, q_s)$. 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 $q_m + q_f = \bar{q} - 1$. They face four options: (1) stop childbearing, (2) have another birth without sex selection, (3) have another child with sex selection for a male, and (4) have another child with sex selection for a female. The value of *not* having an additional 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 sex selection for a *male* child 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 sex selection for a *female* child is

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

Parents decide whether to have another birth ($b \in \{0, 1\}$) and, if so, whether to select a male or female child ($s_f, s_m \in \{0, 1\}$). This involves comparing the value of each option. Therefore, the value of having $\{q_m, q_f, q_s\}$ when $q_m + q_f = \bar{q} - 1$ is

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

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 value function if parents have $\bar{q}-1$ children. Going backward, we can similarly solve for the fertility and sex selection decisions if the number of children is $\bar{q} - 2$, $\bar{q} - 3$, . . . , 0.

## 3.3 Household Formation

Individuals enter the marriage market upon reaching adulthood. Denote the measure of young adults of sex $g \in \{m, f\}$ with human capital $h$ by $Q_g^h$. Given that $p_m > 0.5$ and that parents tend to select more sons than daughters, we have

$$Q_f^{h^L} + Q_f^{h^H} < Q_m^{h^L} + Q_m^{h^H}. \tag{20}$$

To replicate the observed pattern of assortative mating, the marriage market is assumed to clear in three stages, following Fernández and Rogerson (2001) and Fernández et al. (2005). First, a fraction of males are randomly excluded from the market, ensuring an equal measure of males and females.[7] Denote the measure of remaining males with human capital $h$ by $\tilde{Q}_m^h$, then we have

$$Q_f^{h^L} + Q_f^{h^H} = \tilde{Q}_m^{h^L} + \tilde{Q}_m^{h^H}. \tag{21}$$

Second, males and females with the same level of education meet, and a proportion, $\psi min\{\tilde{Q}_m^h, Q_f^h\}$ get matched. The parameter $\psi$ governs the degree of assortative mating, with $\psi = 0$ representing random matching. In the third stage, unmatched individuals

[7] The 2015-2016 DHS data suggest that more educated men tend to marry later. However, by age 35, education level has no significant effect on the likelihood of ever marrying.

are paired randomly.

### 3.4 Equilibrium

I focus on a stationary equilibrium, where young households optimize fertility, sex selection, educational investment, and female labor supply, while middle-aged households optimize female labor supply and savings. Given these decisions and the marriage formation protocol, the distributions of individuals and households by education are constant.

To compute a stationary equilibrium, I solve a fixed point problem iteratively. Starting with an initial guess for the lump-sum transfer from discrimination costs, the distribution of households, intergenerational household-type transitions, and transfers from sons and daughters conditional on their household type, I solve household problems and simulate household formation. The resulting decisions are then used to update the initial guess, and this process repeats until convergence.[8]

## 4 Quantitative Analysis

This section presents the calibration procedures and the model fit. One period in the model corresponds to 20 years, and thus, an individual in the model lives for 80 years. Parents are assumed to have at most $\bar{q} = 6$ children. The low level of human capital in the model corresponds to at most primary education in the data, while the high level of human capital corresponds to at least secondary education. According to the 2015-2016 DHS data, individuals aged 21-54 without secondary education have received about 1.6 years of schooling on average, while those with secondary education have received about 10.5 years of schooling. 65.7% of males and 55.3% of females having completed secondary education.

### 4.1 Calibration

I first set some parameters to their data counterparts or borrow them from the literature. I start with the basic parameters, such as the interest rate and wage rates. Then I move to the parameters related to gender discrimination and dowry. Finally, I

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

explain the calibration of old-age support. The parameters are reported in Table 1.

Table 1: Parameters Chosen outside the Model

| | Description | Value | Source |
|---|---|---|---|
| $\bar{q}$ | Maximum number of children | 6 | |
| $r$ | Annual interest rate | 5.83% | World Bank |
| $h^L, h^H$ | Human capital | 1, 1.621 | Indian HDS 2011-2012 |
| $\psi$ | Assortative mating | 0.557 | Indian DHS 2015-2016 |
| $p_m$ | Natural probability of a male birth | 0.5122 | Normal SRB = 1.05 |
| $\tau$ | Fixed time cost of each child | 0.092 | Deventhal et al. (2021) |
| $\pi$ | Price of sex selection | 0.005 | Arnold et al. (2002), Ganatra and Hirve (2002) |
| $\lambda$ | Gender discrimination in the labor market | 0.67 | Deshpande et al. (2018) |
| $\delta$ | Dowry as a fraction of father's annual income | 0.753 | REDS 1999 |
| $\theta_m$, $\kappa_m$ | Transfer from sons (with a son) | 0.207, 0.972 | LASI 2017-2018 |
| $\theta_f$, $\kappa_f$ | Transfer from daughters (without a son) | 0.119, 0.925 | LASI 2017-2018 |

***Basic Parameters***

The annual interest rate is set to 5.83%, the average real interest rate in India in 1978-2020.[9] To determine the wage rates for the two levels of education/human capital, $\{h^L, h^H\}$, I consider the total earnings (wage earnings, farm income, and business income) for males aged 21-60 using data from the 2011-2012 Indian Human Development Survey (HDS). I restrict the sample to those who work more than 200 hours. The results suggest that men with secondary education earn about 1.621 times more than those without secondary education. Therefore, I normalize $h^L$ to 1 and set $h^H$ to 1.621.[10]

The assortative mating parameter, $\psi$, is set to 0.557 based on the 2015-2016 DHS data for married women aged 21-49 and their spouses.[11] In the data, 27% of marriages are between low-education spouses, and 46% are between high-education ones. The proportion of mixed couples is 26% (19% with a low-education wife and a high-

[9]See https://data.worldbank.org/indicator/FR.INR.RINR?location s=IN.

[10]This number is consistent with the results of Agrawal and Agrawal (2019), who use a Mincerian regression to show that one more year of education increases male wages by 5.0%.

[11]Women aged below 21 years are excluded since 79.8% had never married at the survey time.

education husband and 7% vice versa). With 53.3% of females and 65.7% of males having high education, I choose $\psi$ so that given the household formation protocol, the proportion of different marriages matches the data.

Given the natural male-to-female birth ratio of 1.05, the probability of a male birth in the absence of sex selection, $p_m$, is set to 0.5122. The fixed time cost of each child, $\tau$, is borrowed from Delventhal et al. (2021) and set to 0.092, implying that a mother needs to spend 9.2% of her time to take care of each child.

The monetary cost of sex selection, $\pi$, is set to 0.005. In the model, it is assumed that parents can directly control the sex of their children. However, in reality, parents typically combine prenatal sex detection and abortion to achieve the desired sex. This process involves diagnosing the sex of the fetus and, if it is unwanted, aborting the fetus, repeating this until a child of the desired sex is conceived. Therefore, the number of sex tests and abortions follows a geometric distribution. On average, to achieve a male birth, they need to conduct $1 \div p_m = 1.95$ tests and $1 \div p_m - 1 = 0.95$ abortions. For a female birth, they need $1 \div (1 - p_m) = 2.05$ tests and $1 \div (1 - p_m) - 1 = 1.05$ abortions. For simplicity, assume that parents conduct two tests and one abortion regardless of the targeted sex. The cost of sex detection via ultrasound ranged from 500 to 1000 rupees per test in the 1990s (about 10-20 dollars) (Arnold et al., 2002). I take the middle value of 750 rupees. The cost of an abortion in the second trimester, when the sex of a fetus can be precisely detected, was 1661 rupees (about 35 dollars) on average in 1996–1998 (Ganatra and Hirve, 2002). Therefore, the expected cost of sex selection is $750 \times 2 + 1661 \times 1 = 3111$ rupees. As male income without secondary education is normalized to 1 in the model, we also need to normalize this monetary cost. According to the 2011-2012 Indian HDS data, the median annual income of males without secondary education was 30,854 rupees (about 617 dollars).[12] Given that one model period is 20 years, the cost of each sex selection is about $3111 \div (30854 \times 20) = 0.005$ in the model.

[12]The median is a more appropriate statistic than the mean since it is less sensitive to outliers.

***Gender Discrimination and Dowry***

The gender discrimination parameter in labor market, $\lambda$, is set to 0.67, based on 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 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.*[13]

The value of dowry as a fraction of paternal income, $\delta$, is calculated based on the Rural Economic and Demographic Survey (REDS) in 1999. I consider dowry payments made in the 5 years preceding the survey and divide them by the household income in the previous year, which gives a median dowry-to-income ratio of 0.473.[14] The HDS data in 2011-2012 suggest that a man's share of the household income is about 62.8%, implying that the ratio of dowry to the father's income is about $0.473 \div 0.628 = 0.753$. Given that one model period is 20 years, $\delta$ is set to $0.753 \div 20 = 0.0377$.

***Old-age Support***

Finally, to calibrate the parameters governing old-age support, $\{\theta_m,\ \kappa_m,\ \theta_f,\ \kappa_f\}$, I consider intergenerational coresidence in India. Elderly parents typically receive support by living with their children, effectively transferring a fraction of the children's income to themselves. For this reason, the functional forms for old-age transfers were chosen to reflect the observed patterns of intergenerational coresidence.

Let $\phi$ denote the fraction of children's income transferred to parents *during coresidence*, and $R_g$ denote the fraction of time that each son and daughter spend with their elderly parents ($g \in \{m, f\}$). Then, the fraction of income that a child transfers to his/her parents can be expressed as $t_g(q_m, q_f) = \phi R_g(q_m, q_f)$. Dividing both sides by

[13]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.

[14]Household income is only available for the year preceding the survey. Since income can change over time, dowry payments made more than five years prior to the survey are not considered.

$\phi$, we obtain,

$$R_m(q_m, q_f) = t_m(q_m, q_f)/\phi, \tag{22}$$

and,

$$R_f(q_m, q_f) = t_f(q_m, q_f)/\phi. \tag{23}$$

Notice that $R_m$ and $R_f$ also represent the fraction of parental time spent with each son and daughter, respectively.

Substituting the functions of $t_m$ and $t_f$, as given by Equations (8) and (9), into the above equations, we further get

$$R_m(q_m, q_f) = \begin{cases} \rho_m q_m^{-\kappa_m}, & \text{if } q_m > 0, \\ 0, & \text{if } q_m = 0, \end{cases} \tag{24}$$

and,

$$R_f(q_m, q_f) = \begin{cases} 0, & \text{if } q_m > 0, \\ \rho_f q_f^{-\kappa_f}, & \text{if } q_m = 0, \end{cases} \tag{25}$$

where, $\rho_m \equiv \theta_m/\phi$ and $\rho_f \equiv \theta_f/\phi$ represent the time spent coresiding with an only son and only daughter, respectively, while $\kappa_m$ and $\kappa_f$ capture how coresidence time changes with the number of sons and daughters.

Adding up the time that parents spend with each child, we get the total time spent with children:

$$R_c = \begin{cases} \rho_m q_m^{1-\kappa_m}, & \text{if } q_m > 0, \\ \rho_f q_f^{1-\kappa_f}, & \text{if } q_m = 0. \end{cases} \tag{26}$$

Hence, as long as parents have sons, they reside with them, and the total coresidence time increases in the number of sons. Otherwise, they reside with their daughters, with coresidence time increasing in the number of daughters.

Equation (26) implies that given data on coresidence by the number of sons and daughters, we can estimate $\{\rho_m, \rho_f, \kappa_m, \kappa_f\}$, and, given a value of $\phi$, we can recover $\{\theta_m, \theta_f\}$ from $\{\rho_m, \rho_f\}$.

To estimate $\{\rho_m, \rho_f, \kappa_m, \kappa_f\}$, I use data from the Longitudinal Ageing Study in India (LASI) in 2017-2018. We can observe whether parents were living with their children at the time of the survey, which I use as a proxy for the duration of coresidence.

I run two non-linear least square regressions, one for parents having sons and the other for parents having only daughters. The dependent variable is whether a parent was residing with his/her children, and the independent variable is the number of sons or daughters. The sample includes parents aged 61 years and above. The results show that $\rho_m = 0.713$, $\kappa_m = 0.972$, $\rho_f = 0.409$, and $\kappa_f = 0.925$.

Table 2 displays the proportion of parents residing with their children by the number of sons and daughters in the data, along with the corresponding fitted values. My specification captures two important patterns: (1) parents with sons have a much higher probability of coresidence than those with only daughters (see the blue vs. green shaded area); (2) the probability of coresidence increases gradually with the number of sons but is unaffected by the number of daughters (see the blue shaded area).

Table 2: Proportion 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.

There is no direct evidence on the fraction of children's income transferred to their parents during coresidence. Based on the square root equivalence scale for households with four adults (Eckstein et al., 2019), I set $\phi = 0.29$ and then get $\theta_m = \rho_m \phi = 0.207$

and $\theta_f = \rho_f \phi = 0.119$.[15] These values indicate that an only son transfers 20.7% of his household income to his parents, while an only daughter transfers 11.9%. For parents with a son, the total transfers increase by 0.5–1.9% with each additional son but remain unchanged with additional daughters. For parents with a daughter, the total transfers increases by 1.4-5.1% with each additional daughter but increase more substantially by 74% if they have an additional son.

## 4.2 Estimation and Model Fit

We are left with 12 undetermined parameters: 11 parameters governing household preferences ($\{\beta, \nu_1, \nu_2, \mu_\alpha, \sigma_\alpha, \omega, \epsilon, \eta, e_0, \mu_\xi, \sigma_\xi\}$), and one parameter governing the efficiency of human capital production ($z$). 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. Source: Indian DHS 2015-2016. [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. Source: Indian DHS 2015-2016. [20 targets]

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

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

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

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

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

[15] Under the square root equivalence scale, the per-member consumption in a household with $n$ members is equivalent to $1/\sqrt{n}$ of the total household consumption. For a married couple not living with their parents, each spouse's consumption is $1/\sqrt{2}$ of the household consumption. When their parents join the household, increasing the household size to four, each spouse's consumption decreases to $1/\sqrt{4}$. This represents a 29% reduction in each spouse's consumption.

Table 3 shows moments (iii)-(vii) in the data and model. Moments (i) and (ii) are shown in Figures 5 and 6, 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 1.122, which is very close to the moment in the DHS data (1.131).

Table 3: Data and Model: Selected Moments

| Variable | Data | Model |
|---|---|---|
| *Targeted moments* | | |
| Education expense per son (% income) | 5.05 | 5.86 |
| Education expense per daughter (% income) | 4.26 | 3.60 |
| Fertility difference across women's education | 0.93 | 0.83 |
| Men with secondary education (%) | 65.7 | 70.7 |
| Women with secondary education (%) | 53.3 | 53.0 |
| Female labor supply when young | 0.235 | 0.227 |
| Female labor supply when middle-aged | 0.264 | 0.263 |
| Savings rate when middle-aged (%) | 36.7 | 35.4 |
| *Aggregate moments* | | |
| Fertility | 3.05 | 3.12 |
| Sex ratio at birth | 1.131 | 1.122 |

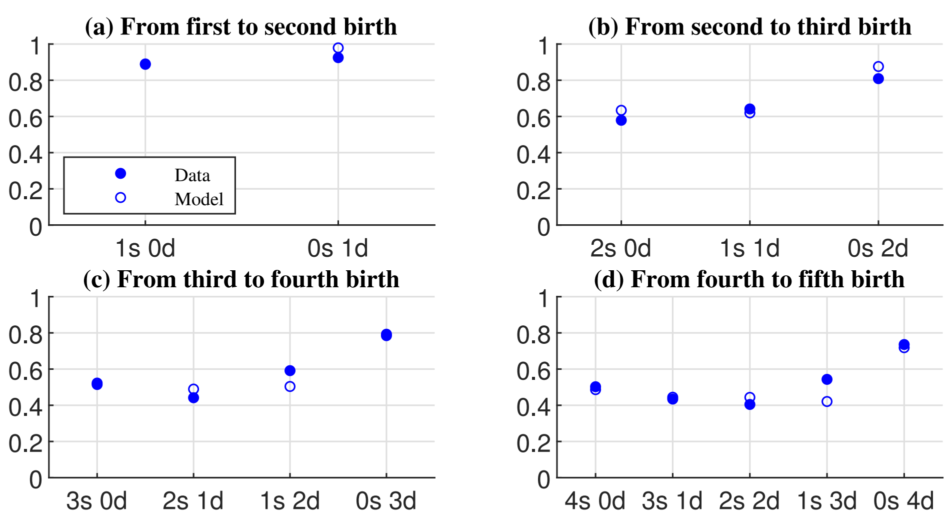



Figure 5: Parity Progression Ratio, Conditional on the Sex Composition of Existing Children (Data and Model).

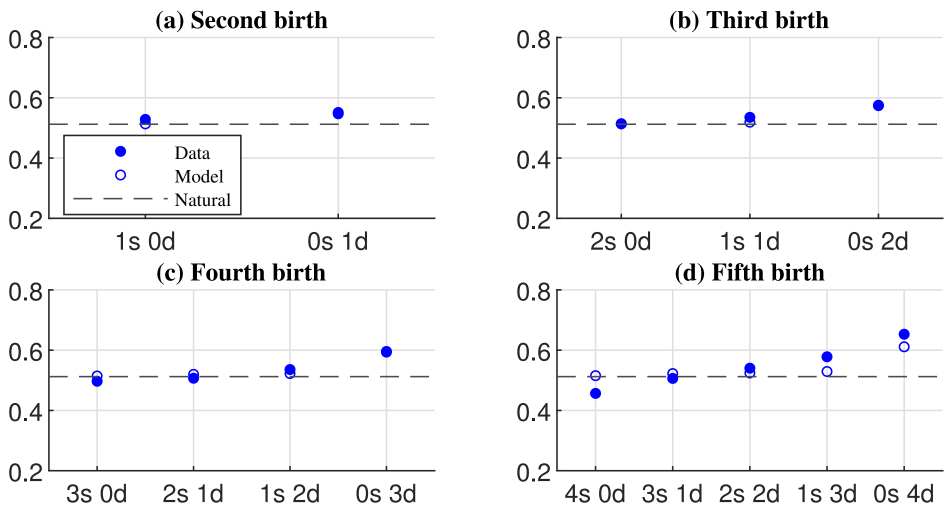



Figure 6: Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children (Data and Model).

In the model, a parameter may affect multiple moments, and a moment may be informative about several parameters. In Appendix B.1, I systematically analyze the relationship between parameter estimates and data moments using the method proposed by Andrews et al. (2017). The following discussion reflects this formal analysis.

The discount factor ($\beta$) primarily affects the savings rate, while $\nu_1$ and $\nu_2$ influence women's labor supply when young and middle-aged, respectively. Additionally, $\nu_1$ impacts fertility decisions due to the time commitment required for each birth. The efficiency of human capital production ($z$) determines educational attainment of men and women, while public education expenditure ($e_0$) affects education expenses per child and fertility decisions. The importance of child quality relative to quantity ($\eta$) influences fertility decisions, and the relative importance of sons ($\omega$) affects education expenses per son compared to daughters. Finally, the distributions of utility weight on children ($\mu_\alpha, \sigma_\alpha$) and disutility from sex selection ($\mu_\xi, \sigma_\xi$), along with the elasticity of substitution between sons and daughters ($\epsilon$), are identified through conditional parity progression ratios (Figure 5) and sex selection decisions (Figure 6).

The estimated parameters are presented in Table 4. The elasticity of substitution between sons and daughters ($\epsilon$) exceeds one, indicating that they are imperfect substitutes. Thus, parents may want another child if the existing children have imbalanced sex composition. In Appendix B.2, I show that when $\epsilon$ is set to 1, i.e., boys and girls are less substitutable, the conditional PPR displays a more pronounced U-shape, and sex selection decisions depend more on the scarcity of sons. The relative weight of sons ($\omega$) is above 0.5, implying an intrinsic preference for sons. The values of $\epsilon$ and $\omega$ suggest that parents are more likely to have another child and select for a male birth if they have daughters but no sons.

Table 4: Estimated Parameters

| | Description | Value | Target |
|---|---|---|---|
| $\beta$ | Annual discount factor | 0.9754 | Moments (vii) |
| $\nu_1$ | Weight of non-market goods when young | 0.2710 | Moments (vi) |
| $\nu_2$ | Weight of non-market goods when middle-aged | 0.5782 | Moments (vi) |
| $\mu_\alpha$, $\sigma_\alpha$ | Distribution of weight of children | -0.9806, 0.4998 | Moments (i)-(v) |
| $\omega$ | Relative importance of sons | 0.5223 | Moments (i)-(v) |
| $\epsilon$ | Substitution between sons and daughters | 35.24 | Moments (i)-(v) |
| $\mu_\xi$, $\sigma_\xi$ | Distribution of disutility from sex selection | -3.5524, 0.2330 | Moments (i)-(v) |
| $\eta$ | Importance of education investment | 0.6093 | Moments (i)-(v) |
| $e_0$ | Public education expenditure | 0.0322 | Moments (i)-(v) |
| $z$ | Efficiency of human capital production | 13.83 | Moments (i)-(v) |

*Notes.* 48 moments are used to estimate 12 parameters.

### 4.3 Non-targeted Moments

In Appendix B.3, I demonstrate that the model replicates well the fertility rates and education expenses per son and daughter *across women's education levels* in the data. It also replicates their sequential fertility behaviors conditional on the sex composition of existing children. Moreover, it matches the distribution of couples by education level.

## 5 External Validation

This section utilizes the model to simulate household responses to changes in dowry values, a conditional transfer program targeted at families with daughters, and a ban on sex selection. The model's predictions are contrasted to empirical findings from quasi-experimental studies in the literature. The comparisons demonstrate that the model can produce predictions consistent with existing empirical evidence.

## 5.1 Changes in the Values of Dowry

In this subsection, I simulate the impact of an increase in dowry value on prenatal sex selection and compare the results with the empirical findings of Bhalotra et al. (2020). To establish causality, they exploit exogenous variation in the price of gold, a key component of dowries in India. Indeed, a 1% increase in the price of gold leads to a 0.783% rise in dowry value (see Table 2 in Bhalotra et al. (2020)).

During 1985-2005, when ultrasound technology was widely available, Bhalotra et al. (2020) find that higher gold prices reduced the likelihood of female births. A 1% increase in gold prices decreased the probability of the second-born child being female by 0.1115 percentage points. When the first child was female, the impact increased to 0.2391 percentage points, reflecting stronger parental incentives to ensure a male child (see Table 5 in Bhalotra et al. (2020)).[16] These findings imply that a 1% increase in dowry value would decrease the probability of the second-born child being female by $0.1115 \div 0.783 = 0.142$ percentage points, and that if the first child is female, the impact is $0.2391 \div 0.783 = 0.305$ percentage points.

In the model, I simulate the effect of dowry value on prenatal sex selection in both short-run partial equilibrium and long-run stationary equilibrium. In the short run, I assume that the distribution of households by human capital, the redistribution of discrimination costs, and parental decision rules remain the same as in the benchmark economy, except for the value of dowry. In the long run, all the variables and decision rules adjust as the economy reaches a new stationary equilibrium. Thus, the short-run effect arises from a sudden change in the dowry value, while the long-run effect stems from a permanent change.

Table 5 presents the results. The model predicts that, in the short run, a 1% increase in the value of dowry would reduce the probability of a second-born child being female by 0.106 percentage points. This reduction is even larger, at 0.190 percentage points, if the first child is female. The long-run effects are similar. Overall, the predicted effects

[16] Since sex selection for the first child is rare, Bhalotra et al. (2020) focus on the second child.

are very close to the estimates of Bhalotra et al. (2020).

Table 5: Effects of An Increase of One Percent in the Value of Dowry

| | Bhalotra et al. (2020) | Model (short run) | Model (long run) |
|---|---|---|---|
| Decrease in the probability that the second child is a girl (pp) | 0.142<br>(0.015, 0.269) | 0.106 | 0.109 |
| Decrease in the probability that the second child is a girl if the first child is a girl (pp) | 0.305<br>(0.148, 0.463) | 0.190 | 0.196 |

*Notes.* 90% confidence intervals in parentheses. Confidence intervals are calculated using the point estimates and standard errors in Table 5 in Bhalotra et al. (2020).

## 5.2 Conditional Cash Transfers

This subsection applies the model to simulate the response of sex selection and fertility behaviors to a conditional cash transfer program named *Dhanlakshmi*, which was implemented in seven Indian states from 2008 to 2013. The model's predictions are compared to the empirical findings of Biswas et al. (2023).

The *Dhanlakshmi* program was initiated in 2008 in 11 blocks across seven states: Andhra Pradesh, Bihar, Chhattisgarh, Jharkhand, Odisha, Punjab, and Uttar Pradesh. Its objectives were to normalize the sex ratio, increase investment in girls' education and healthcare, and delay their marriage. The program is open to girls born on or after November 19, 2008, regardless of their household income or the number of siblings. Parents of eligible girls could receive 5,000 rupees (about 115 dollars) upon presenting the girl's birth registration and an additional 1,250 rupees (about 29 dollars) if she was fully immunized. Moreover, parents could receive 3,500 rupees (about 81 dollars) if their daughter completed primary education (Grade 5) and an additional transfer of 3,750 rupees (about 86 dollars) if she completed secondary education (Grade 8). Finally, if the girl remained unmarried until age 18, she could receive an insurance maturity benefit of 100,000 rupees (about 2,302 dollars). Although the program was discontinued in 2013, its financial commitments continued to be honored.

Biswas et al. (2023) employs a difference-in-difference approach to examine the impact of this program on sex selection (as measured by the probability that a birth is a female) and fertility (as measured by whether a woman has a birth in a given year). 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 female by 5.5 percentage points. In addition, it increased the probability of a woman giving birth in a given year by 0.0085 (see Table 5 in Biswas et al. (2023)).

To simulate the program in the model, I focus on the transfers for female births and education.[17] The transfer for a female birth is normalized to $5000 \div (30854 \times 20) = 0.0081$, adjusting for the median annual income of males without secondary education (30,854 rupees) and a model period of 20 years. Considering that all individuals attain low education in the model, equivalent to 1.6 years of schooling in the data, I assume that parents receive no transfer if their daughter completes low education but receive a total of 7250 rupees if their daughter completes high education. This transfer amounts to $7250 \div (30854 \times 20) = 0.0117$ in the model. Finally, I assume that the transfer for female births is received when parents are young, while the transfer for female education is received in middle age, after their daughters' education achievement is realized.

It should be noted that the sex ratio at birth and fertility rate in Punjab are different from the national averages, with Punjab exhibiting a more skewed sex ratio. To ensure that the simulation results are comparable to the empirical findings of Biswas et al. (2023), the model is recalibrated to the Punjab data, as explained in Appendix C.

The *Dhanlakshmi* program is first simulated in a short-run partial equilibrium, where the distribution of households by human capital, the redistribution of discrimination costs, and parental decision rules, remain unchanged. Next, it is simulated in a long-run stationary equilibrium, where the program is financed by a proportional labor income tax on currently working men and all variables and parental decision rules are

[17] Transfers for immunization and delayed marriage may also affect fertility and sex selection. However, incorporating these measures into the model, which abstracts from health investment and marriage age, is very difficult.

updated.

The results are presented in Table 6. First, the model predicts a 3.7 percentage points in the probability of a female birth both in the short and long run, which closely aligns with the estimate of Biswas et al. (2023). Second, the model predicts an increase in fertility by 0.207 in the short run and 0.198 in the long run, which is consistent with the positive (albeit statistically insignificant) effect of the program on annual fertility risks estimated by Biswas et al. (2023). Overall, the results suggest that the model can deliver reliable predictions in counterfactual analyses.

Table 6: Effects of *Dhanlakshmi*

| | Biswas et al. (2023) | Model (short run) | Model (long run) |
|---|---|---|---|
| Probability of a birth being female before the policy (%) | 45.6 | 43.4 | 43.4 |
| Increase in the probability of a birth being female (pp) | 5.5<br>(0.864, 10.136) | 3.7 | 3.7 |
| Annual fertility risk before the policy | 0.287 | | |
| Increase in annual fertility risk | 0.009<br>(-0.004, 0.022) | | |
| Fertility before the policy | | 2.73 | 2.73 |
| Increase in fertility | | 0.207 | 0.198 |

*Notes*. 90% confidence intervals in parentheses. Confidence intervals are calculated using the point estimates and p-values from the unrestricted wild cluster bootstrap in Table 5 in Biswas et al. (2023).

### 5.3 Sex Selection Bans

In Appendix D, I use the model to simulate the impacts of a sex selection ban. The results suggest that it can reduce the SRB to its natural level by design. Moreover, the increase in female births is concentrated among families with a firstborn girl. These findings are consistent with Nandi and Deolalikar (2013) and Rastogi and Sharma (2022), who exploit the staggered implementation of the sex selection ban across Indian states from the late 1980s to the early 2000s to study its effect on sex ratios.

## 6 Causes and Consequences of Missing Women

How do different factors, including the economic factors and intrinsic son preference, affect the sex ratio? How do they affect the fertility rate, children's education, women's labor supply, and household income? I now use the model economy as a quantitative laboratory to study the causes and consequences of missing women.

## 6.1 Economic Factors

To evaluate how gender differences in economic factors affect the sex ratio at birth, fertility rate, children's education, women's labor supply, and household income, I conduct the following counterfactual experiments:

(i) *Equalize old-age support.* I assume that daughters provide the same old-age support as sons. Transfers from each child, given the total number of children, are then determined by,

$$t(q_m, q_f) = \theta_m (q_m + q_f)^{-\kappa_m}.$$

(ii) *Equalize wages.* I assume no gender discrimination in the labor market, raising the wage ratio of women to men, given the same human capital, $\lambda$, to 1.

(iii) *Equalize marriage payment.* I reduce the dowry payment ($\delta$) by 50% but introduce an equal marriage payment for sons. This ensures gender parity in marriage costs, while leaving the pre-birth expected marriage payment (in the absence of sex selection) unchanged.[18]

(iv) *Equalize all economic factors.* I combine (i), (ii), and (iii), removing all gender differences in the economic factors. The remaining gender asymmetries are that mothers balance childcare, market work, and home production, while fathers work full-time in the market, and parents retain an intrinsic preference for sons.

The results are reported in Table 7.

The results reveal that the impact of economic factors on the sex ratio is not uniform, depending on their influences on the trade-off between sons and daughters, as well as between the quantity and quality of children. In addition, these factors exert varying effects on the fertility rate, children's education, women's labor supply, and household income. Next, I explain the results in detail.

[18] Alternatively, we can eliminate dowry for daughters without introducing marriage payments for sons, reflecting the Dowry Prohibition Act 1961. As shown in Appendix E.3, this also reduces the SRB. However, as the cost of having children decreases, parents now have more children but invest less in their education.

Table 7: Effects of the Gender Difference in Economic Factors

| | Benchmark | Equal support | Equal wages | Equal marriage payment | Equal economic factors |
|---|---|---|---|---|---|
| Sex ratio at birth | 1.122 | 1.071 | 1.220 | 1.050 | 1.050 |
| Fertility | 3.12 | 3.02 | 2.23 | 3.16 | 2.16 |
| Education expense per son (% income) | 5.86 | 5.71 | 8.20 | 5.85 | 8.21 |
| Education expense per daughter (% income) | 3.60 | 4.41 | 5.87 | 3.57 | 6.62 |
| Men with secondary education (%) | 70.7 | 69.9 | 89.7 | 70.5 | 90.0 |
| Women with secondary education (%) | 53.0 | 60.8 | 80.2 | 52.6 | 84.6 |
| Female labor (young) | 0.227 | 0.247 | 0.440 | 0.223 | 0.450 |
| Female labor (middle-aged) | 0.262 | 0.276 | 0.449 | 0.261 | 0.453 |
| Savings rate (middle-aged) | 35.4 | 35.0 | 35.8 | 35.2 | 35.0 |
| Household income | 100.0 | 102.3 | 120.8 | 100.9 | 127.2 |

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

First, if daughters provide equal old-age support as sons, the incentive for parents to abort female fetuses diminishes, reducing the SRB from 1.122 to 1.071. Parents without sons are less inclined to have additional children, slightly lowering the fertility rate by 0.10. The share of household income allocated to each daughter's education increases from 3.60% to 4.41%, raising the proportion of women with secondary education from 53.0% to 60.8%. Although parents slightly reduce spending on sons' education, the impact is minimal. Better education also encourage women to participate more in the labor market, boosting young women's labor supply by 8.8% and middle-aged women's by 5.3%. Consequently, household income grows by 2.3%.

Second, if there is no gender discrimination in the labor market, the SRB would surprisingly increase to 1.220. While higher wages for women may suggest a higher value for daughters, parents typically do not prioritize daughters' income unless it contributes to old-age support. Instead, sons become more valuable, as their wives' higher earnings can be transferred to their parents. Hence, the gender gap in old-age support is widened.[19,20] Moreover, higher wages drive mothers to supply more labor

[19] This aligns with Das Gupta et al. (2003), who show that increasing women's labor supply cannot normalize the sex ratio if parents continue to rely only on sons for old-age support.

[20] If we remove gender disparities in old-age support and wages simultaneously, the SRB will not increase but instead drop to 1.090, confirming their joint effects (see Appendix E.2).

and have fewer children, reducing the fertility rate from 3.12 to 2.23. As fertility declines, the likelihood of having no sons increases, amplifying the incentives for sex selection. [21] These two factors together lead to a more biased sex ratio.

Lower fertility increases educational investment per child. Education spending as a share of household income increases by 2.3 percentage points for both sons and daughters, increasing secondary education attainment by 19.0 percentage points for men and 27.2 for women.[22] Overall, household income increases by 20.8%, due to improved education and female labor supply.

Third, if the marriage costs are equalized between sons and daughters, son-biased sex selection will be eliminated, resulting in a natural SRB of 1.05. However, the effects on other variables, including the fertility rate, education spending, and women's labor supply, are negligible.

Finally, if all gender differences in economic factors are eliminated, the SRB would drop to 1.05. Notably, these factors interact with each other, and eliminating gender discrimination in the labor market no longer increases the sex ratio. Although the fertility rate drops significantly to 2.16, parents now have fewer incentives for son-biased sex selection: daughters provide the same old-age support as sons, and their marriage costs have fallen relative to sons.

The share of household income allocated to each daughter's education increases from 3.60% to 6.62%, raising the proportion of women with secondary education from 53.0% to 84.6%. Sons also benefit from the lower fertility. The share of household income allocated to each son's education increases from 5.86% to 8.21%, raising the proportion of men with secondary education from 70.7% to 90.0%. As women's education improves and the gender pay gap closes, their labor supply surges—98.2% more when young and 72.9% more when middle-aged. Consequently, household income grows by 27.2%.

In conclusion, economic factors affect parental decisions in different ways, with

[21] See Jayachandran (2017) and Ebenstein (2010) for empirical evidence that lower fertility leads to more sex selection.

[22] There is a feedback loop between women's education and fertility: higher education encourages labor market participation and lower fertility, which leads to greater investment in children's education due to the quantity-quality trade-off.

significant interactions among them. Eliminating these gender disparities can normalize the SRB, lower fertility rates, narrow the education gap, and lead to higher female labor supply and substantial income gains. In addition, Figures A.14 and A.15 in Appendix E.1 demonstrate that the practice of son-biased fertility stopping rules and son-biased sex selection is no longer prevalent.

## 6.2 Economic Factors vs. Intrinsic Sex Preference

I now compare the relative importance of gender differences in economic factors and intrinsic son preference. To this end, after equalizing the economic factors for sons and daughters, I further eliminate the intrinsic son preference by setting the son preference parameter ($\omega$) to 0.5. The results are reported in Table 8.

Table 8: Effects of the Gender Difference in Economic Factors vs. Intrinsic Son Preference

| | Benchmark | Equal economic factors | No gender difference | No intrinsic son preference |
|---|---|---|---|---|
| Sex ratio at birth | 1.122 | 1.050 | 1.050 | 1.070 |
| Fertility | 3.12 | 2.16 | 2.14 | 3.09 |
| Education expense per son (% income) | 5.86 | 8.21 | 7.51 | 5.40 |
| Education expense per daughter (% income) | 3.60 | 6.62 | 7.55 | 4.39 |
| Men with secondary education (%) | 70.7 | 90.0 | 88.3 | 67.8 |
| Women with secondary education (%) | 53.0 | 84.6 | 88.5 | 59.9 |
| Female labor (young) | 0.227 | 0.450 | 0.458 | 0.243 |
| Female labor (middle-aged) | 0.262 | 0.453 | 0.461 | 0.281 |
| Household income | 100.0 | 127.2 | 127.7 | 101.4 |

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

The results show that once economic factors are equalized, further eliminating intrinsic sex preference has limited effects on the SRB and gender gap in education. Since the SRB has already normalized, it barely decreases further. The proportion of women with secondary education increases by an additional 3.9 percentage points, but this is much smaller than the increase of 31.6 percentage points from equalizing economic factors. Meanwhile, the proportion of men with secondary education decreases slightly by 1.7 percentage points. In addition, the effects on fertility, women's labor supply, and household income are negligible.

What happens if only intrinsic son preference is removed? The last column in Table 8 shows that the SRB declines and the gender gap in education narrows, but the impacts are smaller compared to equalizing economic factors. Similarly, the effects on fertility, female labor supply, and household income are relatively small.

In conclusion, the imbalanced sex ratio can be largely attributed to gender differences in economic factors. These disparities have also contributed to higher fertility rates, lower education levels, and lower female labor supply.

# 7 Can Policies Normalize the Sex Ratio?

What policies can governments implement to address imbalanced sex ratios? How will they affect fertility rates, children's education, women's labor supply, and household income? To answer these questions, I first compare two widely adopted measures in India: subsidies for female births and female education. I then examine the impacts of a pension system.

## 7.1 Subsidize Female Births vs. Subsidize Female Education

Over several decades, state governments in India have implemented various conditional cash transfer programs to address skewed sex ratios. Two notable initiatives are cash transfers for female births and those tied to daughters completing certain levels of education. Which policy is more effective in normalizing the sex ratio? Which can enhance investment in girls' education? How do they affect fertility, female labor supply, and household income?

This subsection compares the impacts of two policies: a lump-sum subsidy for female births and a proportional subsidy for female education. For the birth subsidy, parents receive a transfer for each daughter, set at 16.2% of the annual income of a male without secondary education, matching the value in the *Dhanlakshmi* program. For the education subsidy, the government funds 14.8% of female education expenses, a rate chosen to achieve the same reduction in the SRB as the birth subsidy. Both subsidies are financed through a proportional labor income tax. The results are presented in

Table 9.

Table 9: Effects of the Subsidies for Female Births and Female Education

| | Benchmark | Birth subsidy | Education Subsidy |
|---|---|---|---|
| Sex ratio at birth | 1.122 | 1.091 | 1.091 |
| Fertility | 3.12 | 3.25 | 2.93 |
| Education expense per son (% income) | 5.86 | 5.61 | 5.74 |
| Education expense per daughter (% income) | 3.60 | 3.41 | 5.14 |
| Men with secondary education (%) | 70.7 | 68.1 | 70.8 |
| Women with secondary education (%) | 53.0 | 50.1 | 66.5 |
| Female labor (young) | 0.227 | 0.213 | 0.262 |
| Female labor (middle-aged) | 0.262 | 0.265 | 0.288 |
| Household income | 100.0 | 98.8 | 103.6 |
| Labor income tax rate (%) | 0.00 | 0.43 | 0.66 |
| Subsidy as a share of GDP (%) | 0.00 | 0.41 | 0.62 |

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

Both policies reduce the SRB to 1.091 by design, but their mechanisms differ. The female birth subsidy directly lowers the cost of having daughters, while the education subsidy reduces the cost of educating them, thereby increasing the utility of having and educating daughters.

Although both policies discourage sex selection, their effects on fertility and education diverge due to their opposite effects on the quantity-quality trade-off. The female birth subsidy reduces the expected cost of childbirth, raising the fertility rate to 3.25 and reducing investment in education. As a result, the proportion of men and women with secondary education declines to 68.1% and 50.1%, respectively. In contrast, the female education subsidy motivates parents to invest more in their daughters' education, increasing the proportion of women with secondary education to 66.5%. While there is a reallocation of resources away from sons, the lower fertility rate (2.93) allows for greater investment per child, leaving men's educational attainment almost unchanged.

The two policies also differ in their impact on women's labor supply through fertility and education. Higher fertility reduces the time spent on market work, while more education increases wages and labor supply. Consequently, the female birth subsidy reduces young women's labor supply by 6.2%, while the female education subsidy increases it by 15.4%. Both subsidies increase the labor supply of middle-aged women, but the effect is minor for the birth subsidy and much more pronounced for the education subsidy. Finally, due to their diverging effects on female education and labor supply, the birth subsidy lowers average household income by 1.2%, while the education subsidy raises it by 3.6%.

The above comparison highlights the overall advantages of the female education subsidy. However, its financial costs are also higher. The female birth subsidy requires a tax rate of 0.43% and expenditures of 0.41% of GDP. In contrast, the female education subsidy requires a tax rate of 0.66% and expenditures of 0.62% of GDP—approximately 50% more.[23]

In conclusion, although the female birth subsidy can reduce the SRB, it has unfavorable effects on fertility, educational attainment, women's labor supply, and household income due to the quantity-quality trade-off. In contrast, a female education subsidy can improve both the SRB and the fertility rate, educational attainment, women's labor supply, and household income, albeit at a higher financial cost.[24]

## 7.2 Pensions

As demonstrated in Section 6, the concern for old-age support is an important driver of son-biased sex selection. However, a pension system can reduce the reliance on children for old-age support, thereby reducing the importance of the gender difference in this regard. How would the pension system affect the SRB? What are its effects on fertility, children's education, women's labor supply, and household income?

To answer these questions, I introduce a pay-as-you-go pension system into the model. I assume that retired men receive pensions, funded by a labor income tax on currently working men. The replacement rate is set to 50%, aligning with the targeted rate of the Employees Provident Fund Scheme, the largest pension system in India (Kim and Bhardwaj, 2011).[25] Given that women supply much less labor, I assume

[23] In Appendix E.4, I choose the subsidy rate for the female education subsidy so that the associated tax rate is the same as for the female birth subsidy. In this case, the SRB is only reduced to 1.104.

[24] The conclusions remain largely unaltered with alternative subsidy rates (see Appendix E.4).

[25] As of the 2011 census, only about 12% of India's workforce was covered by a pension scheme.

that they neither receive pensions nor pay taxes. The results are shown in Table 10.

Table 10: Effects of the Pension System

| | Benchmark | Full support from children | Half support from children | No Support from children |
|---|---|---|---|---|
| Sex ratio at birth | 1.122 | 1.095 | 1.074 | 1.061 |
| Fertility | 3.12 | 3.13 | 3.13 | 3.12 |
| Education expense per son (% income) | 5.86 | 5.78 | 5.50 | 5.15 |
| Education expense per daughter (% income) | 3.60 | 3.37 | 3.57 | 3.79 |
| Men with secondary education (%) | 70.7 | 65.3 | 63.2 | 60.4 |
| Women with secondary education (%) | 53.0 | 46.2 | 48.1 | 49.8 |
| Female labor (young) | 0.227 | 0.262 | 0.265 | 0.269 |
| Female labor (middle-aged) | 0.262 | 0.242 | 0.261 | 0.280 |
| Saving rate when middle-aged (%) | 35.4 | 26.2 | 27.6 | 29.1 |
| Household income | 100.0 | 98.8 | 99.2 | 99.2 |
| Labor income tax rate (%) | 0.00 | 13.42 | 13.22 | 13.10 |

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

The results indicate that a tax rate of 13.42% is required to finance the pension system. As pensions replace the old-age support typically provided by sons, the SRB drops to 1.095. Interestingly, although children become less valuable for old-age support, parents do not reduce their number of children. Instead, they reduce investment in their children's education.[26] Consequently, the proportion of men and women with secondary education drops to 65.3% and 46.2%, respectively.[27] In addition, household income declines by 1.2%.

The above analysis assumes that the pension system does not change the level of old-age support that the parents receive from their children. However, old-age support from children may decrease as a change in social norms over time (Cheng et al., 2018;

[26] The effect of pensions on fertility is ambiguous. On one hand, pensions reduce the benefits of having children, potentially lowering fertility. On the other hand, they may reduce parental investment in children's education, thereby increasing fertility through the quantity-quality trade-off. Empirical studies typically find a negative effect of pension plans on fertility rates (Rossi and Godard, 2022; Danzer and Zyska, 2023). However, I find that while the effect is negative at low replacement rates, it turns positive at high replacement rates (see Appendix E.5).

[27] My findings are consistent with empirical evidence. Bau (2021) show that pension plans reduced investment in the education of sons and daughters in Indonesia and Ghana, respectively, where parents rely on either sons or daughters for old-age support. Ebenstein and Leung (2010) show that the pension plan introduced in rural China in the 1990s reduced son-biased sex selection.

Mukherjee, 2020). To explore this, I consider two hypothetical scenarios where old-aged support is halved or fully eliminated. The SRB now drops further to 1.074 and 1.061, respectively. Moreover, women's educational attainment increases, while men's decreases. This shift occurs because parents reallocate resources from sons to daughters as the old-age support motive for educating sons diminishes

In conclusion, the pension system lowers the SRB but harms the educational attainment of children. If the pension system is accompanied by a reduction in old-age support, the negative effect on girls' education can be alleviated, while the effect on boys' education will be aggravated.

# 8 Conclusions

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. The model focuses on three economic factors that make daughters less valuable for parents: old age support disproportionally provided by sons, the practice of dowry, and gender discrimination against women in the labor market.

The quantitative analysis highlights that economic factors outweigh intrinsic sex preferences in driving sex selection and other household behaviors. In the benchmark model, the SRB is 1.122, with 70.7% of men and 53.0% of women completing secondary education. If gender differences in economic factors are removed, the SRB drops to the biologically normal ratio. The fertility rate declines from 3.1 to 2.2, and the proportion of women with secondary education increases significantly to 84.6%. Men also benefit from the lower fertility rate, with their secondary education attainment rising to 90.0%. In other words, in the benchmark economy, parents' incentives to have sons have resulted in an excess of children, which reduces investment in boys and girls, and more so for girls. Once the economic factors are equalized, the marginal effects of the intrinsic son preference are small.

The model is then used to evaluate two policies aimed at reducing the SRB: a lump-sum subsidy for female births and a proportional subsidy for female education. While both can effectively lower the SRB, they diverge in their broader effects by shifting the parental quantity-quality trade-off in opposite directions. A female birth subsidy at the level of the *Dhanlakshmi* program increases the fertility rate by 0.13, reduces secondary education attainment by 2.9 percentage points for women and 2.6 for men, reduces young women's labor supply by 6.3%, and reduces household income by 1.2%. In contrast, a female education subsidy with a 14.8% rate reduces the fertility rate by 0.20, increases women's secondary education attainment by 13.4 percentage points without affecting men's education, increases young women's labor supply by 15.2%, and increases household income by 3.6%.

Finally, a pension system can reduce the SRB, but it also reduces the educational attainment of children. Indeed, a pay-as-you-go pension system with a replacement rate of 50% can reduce the SRB to 1.095. However, the proportion of men and women with secondary education will decline by 5.4 and 6.8 percentage points, respectively.

## 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.

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.

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.

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

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.

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.*

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, *81*(2, Part 2), S279–S288.

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.

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

Biswas, N., Cornwell, C., & Zimmermann, L. V. (2023). The power of lakshmi: Monetary incentives for raising a girl. *Journal of Human Resources.*

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

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.

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. (2023). Marriage markets and the rise of dowry in India. *Journal of Development Economics*, *164*, 103115.

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.

Danzer, A. M., & Zyska, L. (2023). Pensions and fertility: Microeconomic evidence. *American Economic Journal: Economic Policy*, *15*(2), 126–165.

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.

de la Croix, D., & Licandro, O. (2013). The child is father of the man: Implications for the demographic transition. *The Economic Journal*, *123*(567), 236–261.

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.

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., & 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.

Fernández, 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.

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.

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

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and US economic growth. *Econometrica*, *87*(5), 1439–1474.

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.

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

Li, Q., & Pantano, J. (2023). The demographic consequences of sex-selection technology. *Quantitative Economics*, *14*(1), 309–347.

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

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.

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.

Rossi, P., & Godard, M. (2022). The old-age security motive for fertility: Evidence from the extension of social pensions in Namibia. *American Economic Journal: Economic Policy*, *14*(4), 488–518.

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*, *129*, 1–20.

Soares, R. R. (2005). Mortality reductions, educational attainment, and fertility choice. *American Economic Review*, *95*(3), 580–601.

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

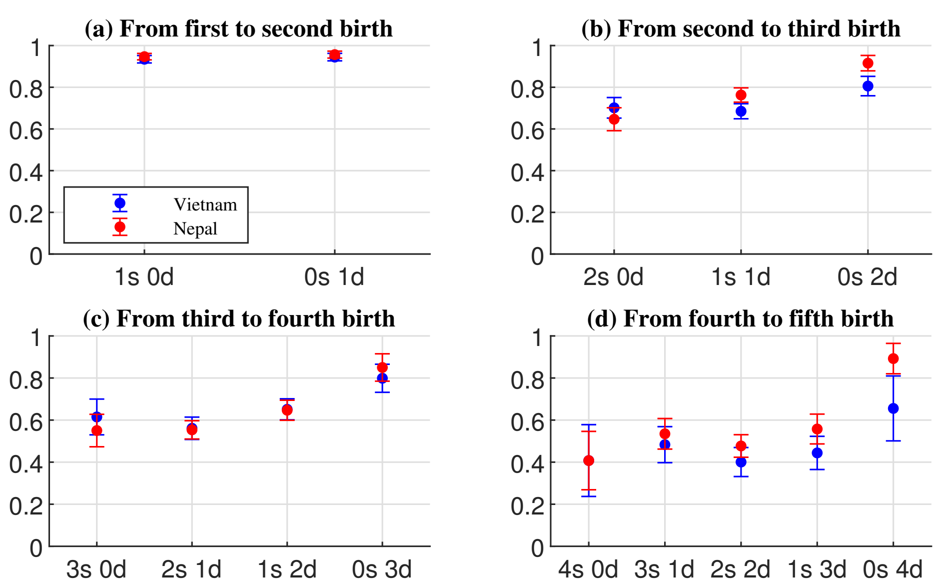
# Missing Women: A Quantitative Analysis

# Online Appendix

## A Additional Empirical Facts

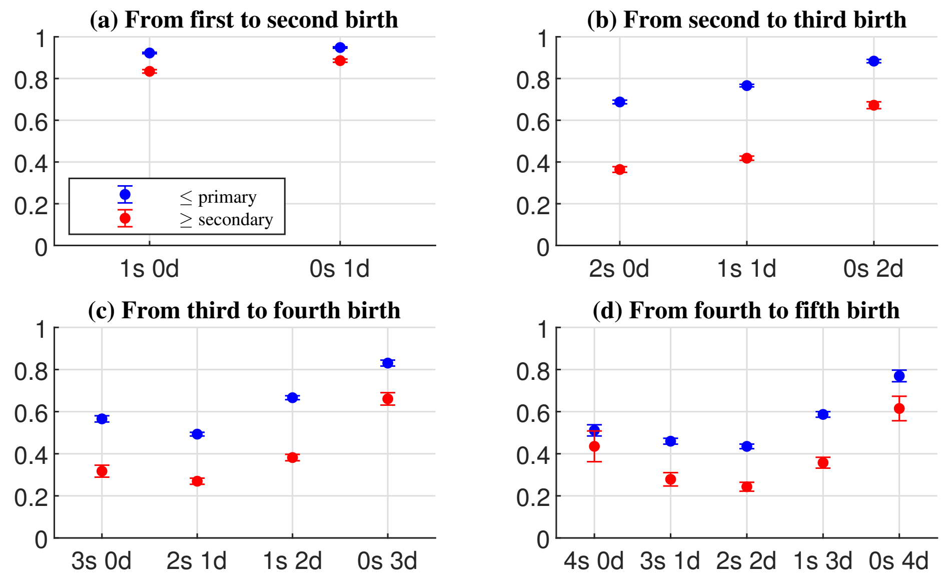
***Robustness Checks for Fact 1: Son-biased Fertility Stopping Rule***

In the main text, this fact is documented using data from the 2015-2016 Indian DHS. For robustness, I first repeat the analysis separately for women with and without secondary education and for women from different castes. The results, shown in Figures A.1 and A.2, indicate that a son-biased fertility stopping rule is prevalent across these groups. Next, I extend the analysis to earlier survey waves from 1992-1993, 1998-1999, and 2005-2006. The results, presented in Figure A.3, reveal that this son-biased fertility stopping rule has persisted over time in India. Finally, I conduct similar analyses for Nepal and Vietnam, where survey data suggest a preference for sons, using data from the 2016 Nepalese DHS and the 2002 Vietnamese DHS. The results, shown in Figure A.4, confirm that parents in these countries also follow a son-biased fertility stopping rule.



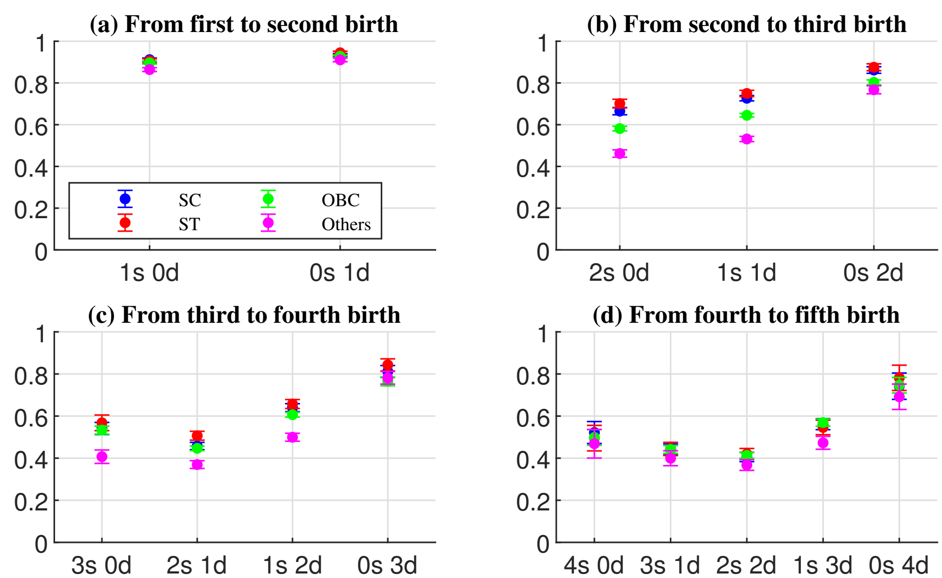

*Notes.* Data are from the Nepalese DHS in 2016 and the Vietnamese DHS in 2002.

Figure A.4: Parity Progression Ratio, Conditional on the Sex Composition of Existing Children, in Nepal and Vietnam.



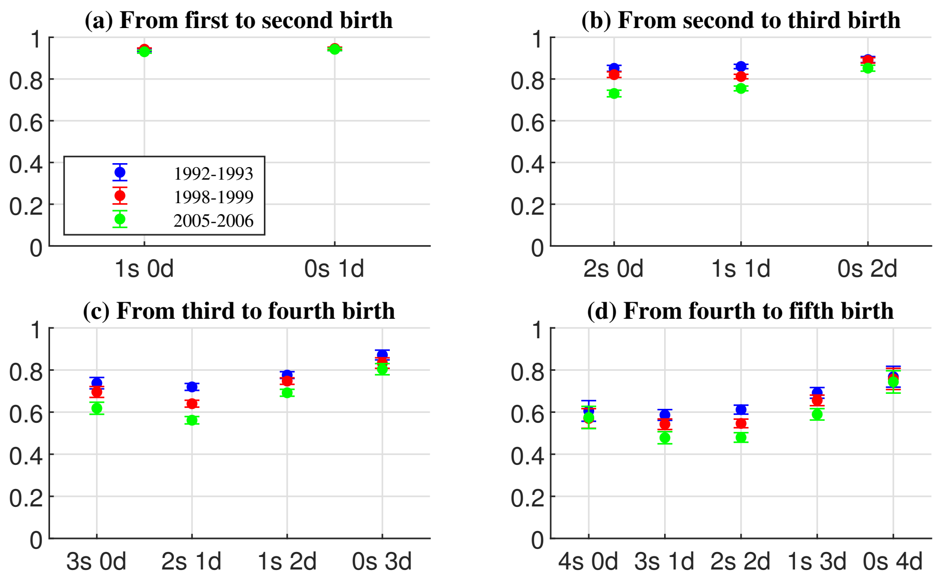

*Notes.* Data are from the Indian DHS in 2015-2016.

Figure A.1: Parity Progression Ratio, Conditional on the Sex Composition of Existing Children, by Women's Education.




*Notes.* Data are from the Indian DHS in 2015-2016. SC: scheduled castes, ST: scheduled tribes, OBC: other backward classes.

Figure A.2: Parity Progression Ratio, Conditional on the Sex Composition of Existing Children, by Women's Caste.



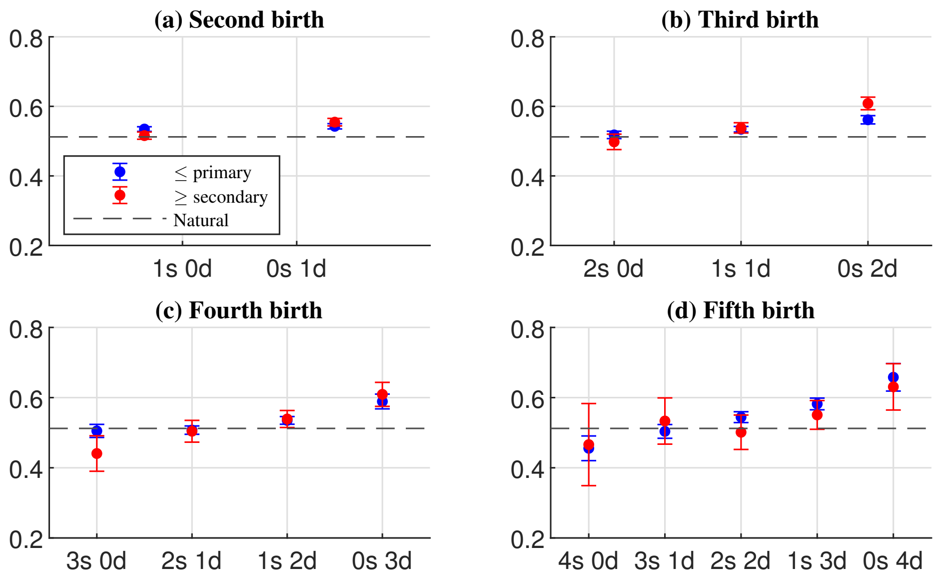

*Notes.* Data are from the Indian DHS in 1992-1993, 1998-1999, and 2005-2006.

Figure A.3: Parity Progression Ratio, Conditional on the Sex Composition of Existing Children, in Previous Surveys.

### *Robustness Checks for Fact 2: Son-biased Sex Selection*

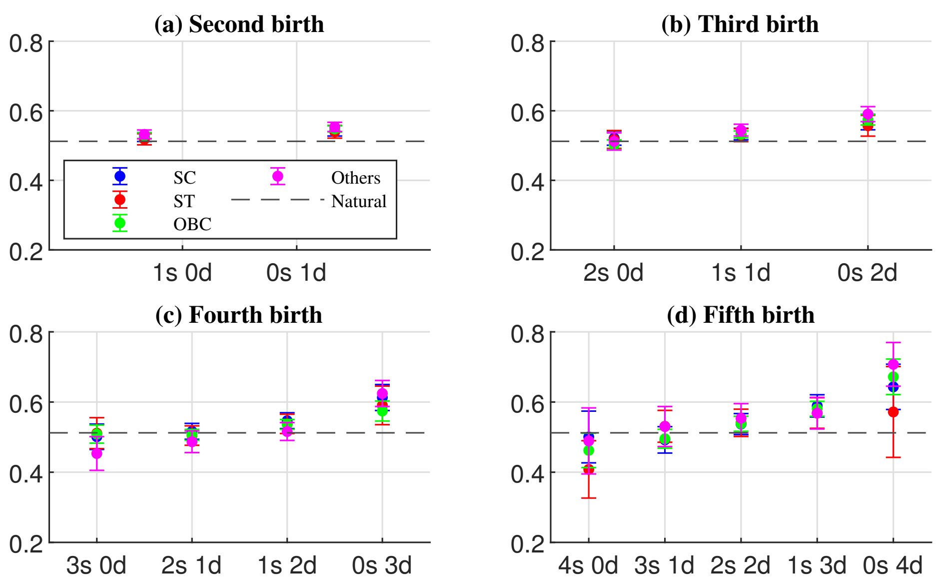
In the main text, this fact is documented using data from the 2015-2016 Indian DHS. For robustness, I first repeat the analysis separately for women with and without secondary education and for women from different castes. The results, shown in Figures A.5 and A.6, indicate that son-biased sex selections are prevalent across these groups.

Next, I extend the analysis to earlier survey waves from 1992-1993, 1998-1999, and 2005-2006. The results, shown in Figure A.7, indicate that sex selection was practiced in earlier periods, but it was less prevalent. This is due to the limited availability of sex selection technology for older cohorts. Ultrasound machines were introduced in India in the mid-1980s and but did not become widely accessible until the mid-1990s.



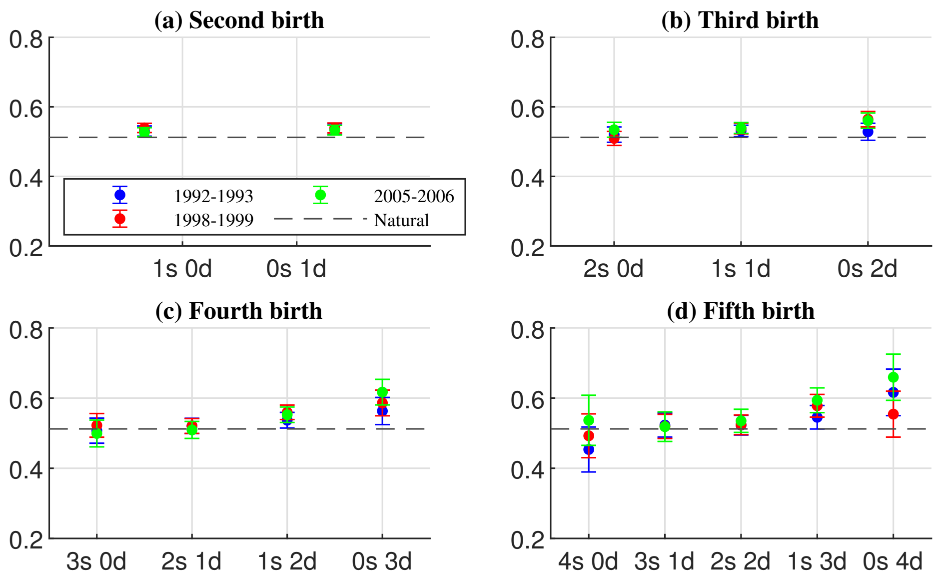

*Notes.* Data are from the Indian DHS in 2015-16.

Figure A.5: 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 2015-2016. SC: scheduled castes, ST: scheduled tribes, OBC: other backward classes.

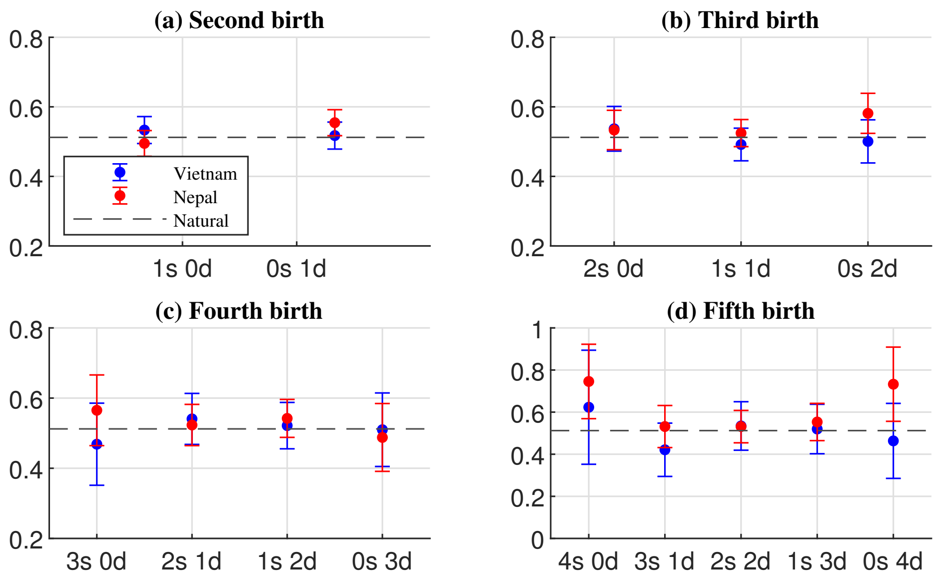
Figure A.6: Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children, by Women's Caste




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

Figure A.7: Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children, in Previous Surveys.

Finally, I conduct similar analyses for Nepal and Vietnam, using data from the 2016 Nepalese DHS and the 2002 Vietnamese DHS. The results, shown in Figure A.8, provide weak evidence that in Nepal, the second and third births were more likely to be boys if the women had no sons. However, there is no consistent pattern for Vietnam. These findings suggest that son-biased sex selection may occur in countries with reported son preferences, but its manifestation depends on both parental desires and the availability of technology.




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

Figure A.8: Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children, in Nepal and Vietnam.

### *Fact 3. Quantity-Quality Trade-off*

To examine the relationship between the quantity and quality of children, I use data from the 2011-2012 Indian Human Development Survey (HDS). The Indian HDS is a nationally representative, longitudinal survey, with the first wave conducted in 2004-2005 and the second wave in 2011-2012. In the second wave, married women reported their own years of education, the number of male and female siblings they had, and their parents' years of education, as well as the same information for their husbands. These data allow for an analysis of the relationship between a wife's or husband's education and the number of children in their natal family.

Using these data, I perform ordinary least square (OLS) regressions separately for married males and females. The dependent variable is the individual's years of schooling, and the key independent variables are the number of male and female children in the natal family. Meanwhile, I control for parental education and the individual's birth year fixed effects.[28] The primary analysis focuses on individuals aged 23-32, with additional age groups used for robustness checks.

[28]The results are robust to additional controls for state, caste, and religion.

The results, presented in the first two columns of Table A.1, suggest a negative correlation between the number of children and their educational attainment. Column (1) shows that an increase in the number of siblings is associated with a decrease in educational attainment for boys. Each additional male sibling is associated with a reduction in schooling of 0.54 years, while each additional female sibling reduces schooling by 0.37 years. Column (2) shows a similar pattern for girls, where each additional male sibling reduces their education by 0.39 years, and each additional female sibling by 0.11 years.

Table A.1: Family Size, Gender and Years of Schooling

| | (1) | (2) | (3) |
|---|---|---|---|
| | Males | Females | 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. $^{***}$ $p < 0.01$.

Interestingly, the impact of siblings on girls' education is less pronounced than on boys'. Moreover, the presence of male siblings has a stronger negative effect on both boys' and girls' education than female siblings. This reveals that educational resources may be disproportionately allocated to boys, a finding consistent with the observation that girls tend to receive less education than boys.

Robustness checks have confirmed these patterns for other age groups (Columns

(1), (2), (4), and (5) of Table A.2) and across caste groups (Columns (1), (2), (4), (5), (7), (8), (10), and (11) of Table A.3).

***Fact 4. Gender Education Gap***

To examine the gender education gap, I use the 2011-2012 Indian HDS data again. I pool the samples of males and females used to document Fact 3, and repeat the regression with an additional dummy variable indicating whether the child is male or female. The results, reported in Column (3) of Table A.1, indicate that female children receive about two fewer years of education than male children, controlling for the number of boys and girls in the family.

Robustness checks have confirmed these patterns for other age groups (Columns (3) and (6) of Table A.2) and across caste groups (Columns (3), (6), (9), and (12) of Table A.3).

Table A.2: 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. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Family Size, Gender and Years of Schooling for Different Castes

| | Scheduled castes | | | Scheduled tribes | | | Other backward classes | | | Others | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | Male | Female | Both | Male | Female | Both | Male | Female | Both | Male | Female | Both |
| Number of male children | -0.425*** | -0.493*** | -0.460*** | -0.472*** | -0.374*** | -0.415*** | -0.630*** | -0.404*** | -0.521*** | -0.508*** | -0.373*** | -0.438*** |
| | (0.094) | (0.103) | (0.071) | (0.135) | (0.129) | (0.096) | (0.061) | (0.075) | (0.049) | (0.087) | (0.095) | (0.069) |
| Number of female children | -0.253*** | -0.088 | -0.184*** | -0.248*** | 0.254 | -0.037 | -0.408*** | -0.155** | -0.312*** | -0.452*** | -0.188** | -0.344*** |
| | (0.073) | (0.097) | (0.059) | (0.094) | (0.159) | (0.088) | (0.052) | (0.063) | (0.041) | (0.072) | (0.093) | (0.061) |
| Male | | | 2.290*** | | | 2.338*** | | | 2.173*** | | | 1.152*** |
| | | | (0.183) | | | (0.288) | | | (0.127) | | | (0.159) |
| Father's years of schooling | 0.326*** | 0.369*** | 0.342*** | 0.445*** | 0.434*** | 0.448*** | 0.317*** | 0.374*** | 0.336*** | 0.344*** | 0.421*** | 0.375*** |
| | (0.030) | (0.033) | (0.022) | (0.062) | (0.054) | (0.042) | (0.019) | (0.020) | (0.014) | (0.023) | (0.024) | (0.017) |
| Mother's years of schooling | 0.525*** | 0.228*** | 0.420*** | 0.367*** | 0.281*** | 0.317*** | 0.491*** | 0.226*** | 0.402*** | 0.287*** | 0.144*** | 0.233*** |
| | (0.042) | (0.051) | (0.034) | (0.089) | (0.082) | (0.060) | (0.024) | (0.026) | (0.018) | (0.025) | (0.028) | (0.019) |
| Birth year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $N$ | 2736 | 2030 | 4766 | 1000 | 752 | 1752 | 5004 | 3438 | 8442 | 3399 | 2180 | 5579 |
| $R^2$ | 0.289 | 0.180 | 0.267 | 0.256 | 0.193 | 0.251 | 0.363 | 0.248 | 0.332 | 0.412 | 0.391 | 0.401 |

*Notes*. Sample consists of men and women aged 23-32 years. Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$.

***Fact 5. Women's Education Gradient in Fertility***

Another important aspect of the quantity-quality trade-off is that more educated parents tend to have fewer children but invest more in their education. To examine the relationship between education and fertility, I use data from the 2015-2016 Indian DHS. I perform an OLS regression where the dependent variable is the number of children ever born, and the key independent variables are the years of education for both the woman and her spouse. Meanwhile, I control for the birth year fixed effects of both spouses.

The results, presented in Table A.4, indicate that higher parental education is correlated with lower fertility, especially for women. Specifically, each additional ten years of education for the husband reduces the number of children by 0.29, while each additional ten years of education for the wife reduces the number of children by 1.17. Robustness checks suggest that the relationships hold across caste groups (Table A.5) and over time when analyzed using previous waves of the survey from 1992-1993, 1998-1999, and 2005-2006 (Table A.6).

Table A.4: Education and Fertility

| | Children ever born |
|---|---|
| Wife's years of schooling | -0.117*** |
| | (0.003) |
| Husband's years of schooling | -0.029*** |
| | (0.003) |
| Wife's birth year fixed effects | Yes |
| Husband's birth year fixed effects | Yes |
| $N$ | 23023 |
| $R^2$ | 0.151 |

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

Table A.5: Education and Fertility for Women from Different Castes

| | (1) Scheduled castes | (2) Scheduled tribes | (3) Other backward classes | (4) Others |
|---|---|---|---|---|
| Wife's years of schooling | -0.126*** | -0.066*** | -0.128*** | -0.114*** |
| | (0.009) | (0.009) | (0.005) | (0.005) |
| Husband's years of schooling | -0.021** | -0.029*** | -0.025*** | -0.027*** |
| | (0.007) | (0.008) | (0.004) | (0.006) |
| Wife's birth year fixed effects | Yes | Yes | Yes | Yes |
| Husband's birth year fixed effects | Yes | Yes | Yes | Yes |
| $N$ | 3765 | 4025 | 8925 | 5156 |
| $R^2$ | 0.141 | 0.063 | 0.165 | 0.206 |

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

Table A.6: Education and Fertility in Previous Surveys

| | (1) India 1992-93 | (2) India 1998-99 | (3) India 2005-06 |
|---|---|---|---|
| Wife's years of schooling | -0.161*** | -0.162*** | -0.147*** |
| | (0.006) | (0.005) | (0.003) |
| Husband's years of schooling | -0.044*** | -0.036*** | -0.030*** |
| | (0.005) | (0.004) | (0.003) |
| Wife's birth year fixed effects | Yes | Yes | Yes |
| Husband's birth year fixed effects | Yes | Yes | Yes |
| $N$ | 16199 | 17025 | 22392 |
| $R^2$ | 0.144 | 0.174 | 0.175 |

*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 the sensitivity of parameter estimates to data moments using the measures from Andrews et al. (2017). Table A.7 presents the percentage change in moments when a parameter changes by one percent, while Table A.8 presents the percentage change in parameters when a moment changes by one percent.

Table A.7: Sensitivity of Moments to Parameters

| | $\beta$ | $\nu_1$ | $\nu_2$ | $\mu_\alpha$ | $\sigma_\alpha$ | $\omega$ | $\epsilon$ | $\mu_\xi$ | $\sigma_\xi$ | $\eta$ | $e_0$ | $z$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Savings rate (middle-aged) | **0.75** | 0.00 | 0.00 | -0.01 | 0.00 | 0.14 | -0.00 | -0.08 | 0.01 | 0.03 | -0.01 | 0.00 |
| Female labor supply (young) | 0.51 | **-1.41** | 0.15 | -0.39 | 0.00 | -1.66 | 0.01 | -0.25 | 0.01 | 2.19 | -0.79 | **0.20** |
| Female labor supply (middle-aged) | **0.72** | 0.08 | **-1.90** | 0.14 | 0.03 | -1.69 | -0.00 | -0.04 | 0.00 | -0.11 | -0.06 | 0.04 |
| PPR if 1s 0d | -0.17 | -0.02 | -0.03 | 0.29 | -0.21 | -0.51 | -0.01 | 0.14 | -0.01 | -0.58 | 0.15 | -0.02 |
| PPR if 0s 1d | 0.01 | -0.00 | -0.00 | 0.06 | -0.06 | 0.20 | -0.00 | -0.05 | 0.00 | -0.08 | 0.02 | -0.01 |
| PPR if 2s 0d | -0.53 | -0.12 | -0.04 | 0.74 | -0.03 | -1.22 | 0.01 | 0.10 | -0.01 | -1.62 | 0.57 | -0.12 |
| PPR if 1s 1d | -0.62 | -0.09 | -0.11 | **0.82** | -0.04 | 0.71 | 0.01 | 0.32 | -0.01 | -1.86 | 0.65 | -0.14 |
| PPR if 0s 2d | 0.06 | -0.01 | 0.02 | 0.28 | -0.06 | 1.19 | -0.01 | -0.38 | 0.01 | -0.43 | 0.13 | -0.05 |
| PPR if 3s 0d | -0.09 | 0.04 | -0.07 | 0.77 | 0.30 | -1.78 | **-0.11** | 0.12 | -0.01 | -2.04 | 0.22 | -0.17 |
| PPR if 2s 1d | -0.66 | -0.13 | **-0.18** | 0.71 | **0.74** | -0.16 | 0.00 | 0.20 | -0.01 | -2.21 | 0.77 | -0.14 |
| PPR if 1s 2d | **-0.79** | -0.07 | -0.06 | **0.85** | 0.09 | 2.16 | 0.01 | 0.10 | 0.02 | **-2.40** | 0.71 | **-0.19** |
| PPR if 0s 3d | 0.04 | -0.22 | -0.01 | 0.41 | -0.04 | 2.24 | -0.04 | -0.63 | 0.01 | -0.82 | 0.44 | -0.10 |
| PPR if 4s 0d | **-0.85** | **-0.84** | -0.01 | 0.75 | 0.61 | -1.03 | 0.01 | 0.08 | -0.01 | -1.52 | **1.35** | -0.13 |
| PPR if 3s 1d | -0.08 | -0.44 | **-0.44** | 0.41 | **1.30** | 0.05 | 0.03 | 0.13 | -0.01 | -1.72 | 0.49 | -0.17 |
| PPR if 2s 2d | **-0.82** | **-0.68** | 0.15 | **0.91** | 0.46 | 0.66 | 0.00 | -0.10 | 0.00 | -1.67 | 0.33 | -0.16 |
| PPR if 1s 3d | **-0.90** | **-1.34** | **-0.29** | **1.03** | **1.00** | **2.30** | 0.01 | -0.06 | 0.03 | -1.93 | 0.34 | -0.18 |
| PPR if 0s 4d | 0.05 | -0.16 | 0.01 | 0.51 | 0.11 | 2.19 | **-0.18** | -0.69 | -0.01 | -0.83 | 0.31 | 0.01 |
| PPR if 5s 0d | -0.63 | **-0.67** | **-0.47** | 0.25 | **1.50** | -0.25 | **-0.46** | 0.11 | **0.25** | **-2.57** | 0.59 | **-0.50** |
| PPR if 4s 1d | **-0.81** | **-1.29** | 0.15 | **0.82** | 0.33 | -0.37 | **-0.32** | 0.15 | -0.00 | **-2.67** | **1.29** | 0.11 |
| PPR if 3s 2d | 0.56 | **-0.49** | **0.22** | **1.04** | **1.43** | -0.82 | -0.00 | -0.08 | 0.01 | -2.01 | **0.85** | **-0.20** |
| PPR if 2s 3d | -0.17 | 0.18 | **-0.41** | **0.83** | **1.30** | -0.09 | -0.00 | -0.06 | -0.01 | **-2.40** | **1.55** | **-0.22** |
| PPR if 1s 4d | -0.23 | -0.02 | 0.09 | **0.87** | **1.28** | 1.96 | **-0.20** | -0.53 | 0.04 | **-3.02** | **1.54** | -0.02 |
| PPR if 0s 5d | 0.12 | -0.37 | 0.02 | 0.51 | **0.66** | 1.66 | **-0.06** | -0.31 | -0.02 | -1.29 | 0.48 | -0.13 |
| Fertility difference across women's edu | **-0.67** | **-0.72** | **-0.34** | 0.16 | -0.29 | 0.65 | -0.02 | 0.11 | 0.01 | 0.04 | 0.55 | 0.02 |
| Men with secondary edu (%) | 0.53 | -0.02 | 0.06 | 0.21 | 0.08 | **-3.41** | 0.01 | -0.03 | -0.00 | **2.77** | **-1.07** | **0.96** |
| Women with secondary edu (%) | 0.39 | -0.02 | 0.04 | 0.06 | 0.03 | 0.78 | 0.01 | -0.03 | -0.00 | 1.94 | -0.60 | **0.67** |
| Edu expense per son | 0.58 | 0.09 | 0.04 | 0.12 | 0.07 | 1.54 | 0.02 | 0.02 | -0.01 | **3.14** | **-0.88** | 0.16 |
| Edu expense per daughter | 0.64 | 0.09 | 0.05 | 0.31 | 0.14 | **-5.23** | 0.01 | -0.04 | -0.00 | **3.55** | **-1.34** | **0.32** |
| Prob that next child is a son if 1s 0d | 0.03 | 0.00 | 0.01 | -0.00 | 0.00 | 0.77 | 0.00 | -0.13 | 0.01 | 0.01 | -0.00 | 0.00 |
| Prob that next child is a son if 0s 1d | 0.38 | 0.03 | 0.08 | -0.09 | 0.02 | 1.19 | -0.01 | **-1.02** | 0.07 | 0.36 | -0.11 | 0.03 |
| Prob that next child is a son if 2s 0d | 0.04 | 0.00 | 0.01 | -0.00 | -0.00 | 1.20 | 0.01 | -0.17 | 0.02 | 0.00 | -0.00 | -0.00 |
| Prob that next child is a son if 1s 1d | 0.14 | 0.00 | 0.03 | -0.00 | -0.00 | 1.16 | 0.00 | -0.52 | 0.06 | 0.06 | -0.02 | 0.00 |
| Prob that next child is a son if 0s 2d | 0.53 | 0.04 | 0.11 | -0.08 | -0.03 | 1.09 | -0.04 | **-1.21** | 0.05 | 0.47 | -0.16 | 0.05 |
| Prob that next child is a son if 3s 0d | 0.06 | 0.00 | 0.01 | -0.00 | -0.00 | 2.01 | 0.01 | -0.25 | 0.03 | 0.00 | 0.00 | -0.00 |
| Prob that next child is a son if 2s 1d | 0.17 | 0.01 | 0.04 | 0.00 | -0.00 | 1.56 | 0.01 | -0.61 | 0.06 | 0.05 | -0.02 | 0.00 |
| Prob that next child is a son if 1s 2d | 0.18 | 0.01 | 0.04 | 0.00 | 0.01 | 1.02 | -0.01 | -0.63 | 0.05 | 0.10 | -0.03 | 0.01 |
| Prob that next child is a son if 0s 3d | 0.57 | 0.15 | 0.13 | -0.06 | -0.03 | 1.02 | **-0.06** | **-1.16** | 0.03 | 0.48 | -0.20 | 0.07 |
| Prob that next child is a son if 4s 0d | 0.10 | 0.00 | 0.02 | -0.00 | -0.00 | **3.36** | 0.02 | -0.39 | 0.04 | -0.01 | 0.00 | -0.00 |
| Prob that next child is a son if 3s 1d | 0.23 | 0.01 | 0.06 | 0.01 | -0.00 | **2.29** | 0.02 | -0.83 | **0.09** | 0.04 | -0.01 | 0.00 |
| Prob that next child is a son if 2s 2d | 0.23 | 0.01 | 0.05 | 0.01 | 0.02 | 1.31 | -0.00 | -0.73 | **0.07** | 0.08 | -0.02 | 0.01 |
| Prob that next child is a son if 1s 3d | 0.26 | 0.04 | 0.05 | 0.01 | 0.02 | 1.30 | -0.02 | -0.85 | **0.07** | 0.13 | -0.02 | 0.02 |
| Prob that next child is a son if 0s 4d | 0.38 | 0.02 | 0.07 | -0.04 | -0.01 | 0.76 | -0.04 | **-1.00** | -0.01 | 0.34 | -0.16 | 0.05 |
| Prob that next child is a son if 5s 0d | 0.12 | -0.01 | 0.03 | -0.00 | -0.00 | **5.39** | 0.03 | -0.57 | 0.06 | -0.00 | -0.00 | 0.00 |
| Prob that next child is a son if 4s 1d | 0.39 | 0.01 | 0.08 | 0.02 | 0.04 | **3.70** | 0.04 | **-1.27** | **0.14** | 0.08 | -0.04 | 0.00 |
| Prob that next child is a son if 3s 2d | 0.29 | 0.00 | 0.06 | 0.02 | 0.05 | 2.06 | 0.01 | **-1.00** | **0.10** | 0.11 | -0.05 | 0.02 |
| Prob that next child is a son if 2s 3d | 0.32 | -0.00 | 0.08 | 0.03 | 0.06 | 1.74 | -0.03 | -0.97 | **0.07** | 0.17 | -0.07 | 0.02 |
| Prob that next child is a son if 2s 3d | 0.36 | 0.01 | 0.07 | 0.06 | 0.09 | 1.56 | -0.03 | **-1.12** | 0.04 | 0.32 | -0.12 | 0.03 |
| Prob that next child is a son if 0s 5d | 0.27 | -0.03 | 0.04 | 0.15 | 0.18 | **2.97** | **-0.08** | **-2.53** | **-0.16** | 0.00 | 0.02 | -0.00 |

*Notes.* Moments most sensitive to each parameter are in bold.

Table A.8: Sensitivity of Parameters to Moments

| | $\beta$ | $\nu_1$ | $\nu_2$ | $\mu_\alpha$ | $\sigma_\alpha$ | $\omega$ | $\epsilon$ | $\mu_\xi$ | $\sigma_\xi$ | $\eta$ | $e_0$ | $z$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Savings rate (middle-aged) | **0.23** | -0.07 | 0.03 | 0.03 | **-0.07** | 0.01 | -0.05 | 0.07 | -0.01 | -0.03 | 0.02 | -0.04 |
| Female labor supply (young) | 0.11 | **-0.29** | 0.02 | -0.09 | -0.01 | -0.01 | **0.33** | -0.02 | 0.04 | -0.00 | 0.00 | -0.08 |
| Female labor supply (middle-aged) | **0.13** | -0.06 | **-0.41** | 0.05 | -0.11 | -0.00 | 0.01 | -0.00 | -0.03 | -0.03 | -0.01 | 0.05 |
| PPR if 1s 0d | 0.03 | 0.00 | -0.00 | **0.09** | -0.08 | -0.00 | -0.05 | 0.01 | **0.11** | -0.01 | -0.01 | -0.03 |
| PPR if 0s 1d | 0.01 | -0.00 | -0.00 | **0.02** | -0.02 | 0.00 | -0.01 | 0.00 | **0.04** | -0.00 | -0.00 | -0.01 |
| PPR if 2s 0d | -0.03 | 0.03 | 0.00 | 0.16 | -0.12 | -0.02 | **0.18** | -0.05 | **0.44** | 0.00 | 0.01 | -0.07 |
| PPR if 1s 1d | -0.00 | 0.03 | -0.02 | **0.17** | -0.10 | 0.01 | 0.01 | 0.03 | **0.16** | 0.00 | -0.01 | -0.05 |
| PPR if 0s 2d | 0.05 | 0.00 | 0.00 | **0.09** | -0.07 | 0.00 | -0.05 | 0.00 | **0.10** | 0.00 | -0.00 | -0.07 |
| PPR if 3s 0d | 0.10 | 0.05 | 0.03 | 0.16 | -0.08 | -0.01 | **-0.26** | 0.01 | -0.08 | -0.09 | **-0.20** | -0.06 |
| PPR if 2s 1d | -0.13 | 0.06 | -0.03 | -0.01 | 0.10 | -0.00 | **0.28** | -0.04 | -0.04 | -0.03 | -0.07 | **0.14** |
| PPR if 1s 2d | -0.05 | 0.05 | -0.02 | **0.12** | -0.05 | 0.01 | 0.09 | 0.01 | **0.17** | -0.03 | -0.07 | 0.03 |
| PPR if 0s 3d | 0.08 | -0.03 | -0.01 | 0.10 | -0.10 | 0.01 | **-0.13** | 0.01 | 0.03 | 0.02 | 0.07 | **-0.11** |
| PPR if 4s 0d | -0.11 | -0.07 | -0.00 | 0.02 | -0.01 | -0.02 | **0.35** | -0.07 | **0.38** | 0.12 | 0.28 | -0.11 |
| PPR if 3s 1d | -0.01 | -0.07 | -0.07 | -0.15 | 0.19 | 0.01 | **0.46** | -0.00 | **-0.41** | -0.07 | -0.13 | 0.13 |
| PPR if 2s 2d | -0.14 | -0.05 | 0.03 | 0.11 | 0.02 | -0.00 | **0.32** | -0.05 | 0.02 | -0.02 | **-0.18** | -0.07 |
| PPR if 1s 3d | -0.10 | -0.13 | -0.07 | 0.02 | 0.08 | 0.02 | **0.40** | -0.01 | -0.14 | -0.03 | **-0.18** | 0.01 |
| PPR if 0s 4d | 0.06 | 0.03 | -0.01 | 0.11 | -0.06 | 0.02 | **-0.65** | 0.03 | **-0.43** | 0.01 | -0.01 | -0.03 |
| PPR if 5s 0d | -0.12 | 0.05 | -0.05 | -0.06 | 0.12 | -0.00 | **-0.83** | -0.00 | **0.76** | 0.02 | -0.09 | -0.17 |
| PPR if 4s 1d | 0.09 | -0.14 | 0.04 | -0.00 | -0.13 | 0.00 | **-0.72** | 0.05 | **-0.26** | -0.02 | 0.12 | 0.25 |
| PPR if 3s 2d | **0.30** | -0.10 | 0.15 | 0.09 | 0.03 | -0.01 | **0.29** | 0.07 | 0.04 | -0.03 | 0.02 | -0.14 |
| PPR if 2s 3d | -0.03 | 0.13 | -0.06 | 0.01 | 0.11 | -0.01 | **0.16** | -0.02 | 0.08 | 0.10 | **0.28** | -0.02 |
| PPR if 1s 4d | 0.00 | 0.10 | 0.05 | -0.03 | 0.08 | 0.01 | **-0.29** | 0.01 | -0.03 | 0.00 | 0.10 | **0.23** |
| PPR if 0s 5d | 0.12 | -0.06 | 0.03 | 0.03 | 0.03 | 0.02 | **-0.19** | 0.05 | **-0.60** | -0.02 | -0.04 | -0.06 |
| Fertility difference across women's edu | -0.09 | -0.08 | -0.11 | 0.03 | -0.08 | 0.00 | 0.01 | -0.02 | **0.23** | 0.10 | **0.20** | -0.07 |
| Men with secondary edu (%) | -0.07 | 0.05 | -0.01 | 0.01 | 0.08 | -0.02 | 0.14 | -0.04 | **0.39** | -0.05 | -0.09 | **0.63** |
| Women with secondary edu (%) | -0.02 | 0.03 | -0.02 | -0.00 | **0.07** | 0.02 | -0.07 | 0.05 | -0.03 | -0.02 | -0.03 | **0.49** |
| Edu expense per son | 0.02 | 0.05 | -0.02 | 0.28 | 0.03 | 0.04 | **-0.63** | 0.11 | -0.40 | 0.19 | 0.08 | **-0.42** |
| Edu expense per daughter | -0.06 | 0.06 | 0.02 | 0.22 | 0.01 | -0.03 | -0.15 | -0.05 | **0.31** | 0.10 | -0.01 | **-0.26** |
| Prob that next child is a son if 1s 0d | 0.00 | -0.00 | -0.00 | 0.00 | 0.00 | 0.00 | **0.02** | 0.00 | **0.04** | -0.00 | -0.00 | 0.01 |
| Prob that next child is a son if 0s 1d | 0.00 | -0.01 | 0.01 | -0.01 | -0.01 | -0.01 | **0.15** | -0.06 | **0.44** | -0.00 | 0.02 | 0.01 |
| Prob that next child is a son if 2s 0d | 0.00 | -0.00 | -0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | **0.04** | -0.00 | -0.00 | **0.01** |
| Prob that next child is a son if 1s 1d | -0.00 | -0.00 | 0.00 | 0.00 | -0.01 | -0.00 | **0.13** | -0.03 | **0.38** | -0.00 | 0.01 | 0.02 |
| Prob that next child is a son if 0s 2d | 0.03 | -0.01 | 0.02 | 0.00 | -0.03 | -0.01 | 0.02 | **-0.07** | **0.28** | -0.00 | 0.02 | -0.02 |
| Prob that next child is a son if 3s 0d | 0.01 | -0.00 | -0.01 | 0.00 | 0.00 | 0.01 | 0.01 | 0.02 | **0.04** | -0.00 | -0.01 | **0.02** |
| Prob that next child is a son if 2s 1d | 0.00 | -0.00 | 0.00 | 0.01 | -0.02 | -0.00 | **0.16** | -0.03 | **0.45** | -0.00 | 0.01 | 0.02 |
| Prob that next child is a son if 1s 2d | -0.00 | 0.00 | 0.00 | 0.01 | -0.02 | -0.01 | **0.10** | -0.04 | **0.36** | 0.00 | 0.02 | 0.01 |
| Prob that next child is a son if 0s 3d | 0.04 | 0.02 | 0.03 | 0.01 | -0.02 | -0.01 | **-0.14** | -0.05 | **0.07** | -0.01 | 0.00 | -0.01 |
| Prob that next child is a son if 4s 0d | 0.02 | -0.01 | -0.01 | 0.01 | 0.00 | 0.03 | 0.01 | 0.04 | **0.06** | -0.00 | -0.01 | **0.04** |
| Prob that next child is a son if 3s 1d | 0.00 | -0.01 | 0.01 | 0.01 | -0.02 | -0.00 | **0.24** | -0.04 | **0.63** | -0.00 | 0.02 | 0.04 |
| Prob that next child is a son if 2s 2d | 0.00 | -0.00 | 0.01 | 0.01 | -0.02 | -0.01 | **0.15** | -0.04 | **0.47** | -0.00 | 0.02 | 0.02 |
| Prob that next child is a son if 1s 3d | -0.00 | 0.01 | 0.01 | 0.01 | -0.02 | -0.01 | **0.08** | -0.04 | **0.42** | 0.00 | 0.02 | 0.01 |
| Prob that next child is a son if 0s 4d | 0.01 | -0.00 | 0.00 | -0.01 | -0.00 | -0.01 | **-0.06** | -0.06 | **-0.18** | -0.01 | -0.02 | -0.00 |
| Prob that next child is a son if 5s 0d | 0.03 | -0.01 | -0.03 | 0.01 | 0.01 | 0.05 | 0.03 | **0.07** | 0.06 | -0.01 | -0.03 | **0.09** |
| Prob that next child is a son if 4s 1d | 0.01 | -0.02 | 0.01 | 0.01 | -0.03 | -0.00 | **0.42** | -0.05 | **0.96** | -0.01 | 0.02 | 0.06 |
| Prob that next child is a son if 3s 2d | -0.00 | -0.00 | 0.01 | 0.01 | -0.02 | -0.01 | **0.24** | -0.05 | **0.68** | -0.01 | 0.02 | 0.04 |
| Prob that next child is a son if 2s 3d | 0.01 | 0.00 | 0.01 | 0.01 | -0.02 | -0.00 | **0.07** | -0.04 | **0.37** | -0.00 | 0.01 | 0.01 |
| Prob that next child is a son if 1s 4d | -0.00 | 0.01 | 0.01 | 0.02 | -0.01 | -0.01 | 0.00 | **-0.05** | **0.18** | 0.01 | 0.00 | -0.02 |
| Prob that next child is a son if 0s 5d | -0.14 | 0.03 | -0.06 | -0.07 | 0.10 | -0.00 | **-0.35** | -0.19 | **-1.70** | 0.01 | -0.08 | -0.08 |

*Notes.* Parameters most sensitive to each moment are in bold.

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

If the substitution elasticity between sons and daughters ($\epsilon$) is set to 1, i.e., they become less substitutable, the PPR conditional on the sex composition of existing children displays a more pronounced U-shape (Figure A.9), and sex selection decisions depend more on the scarcity of sons (Figure A.10).

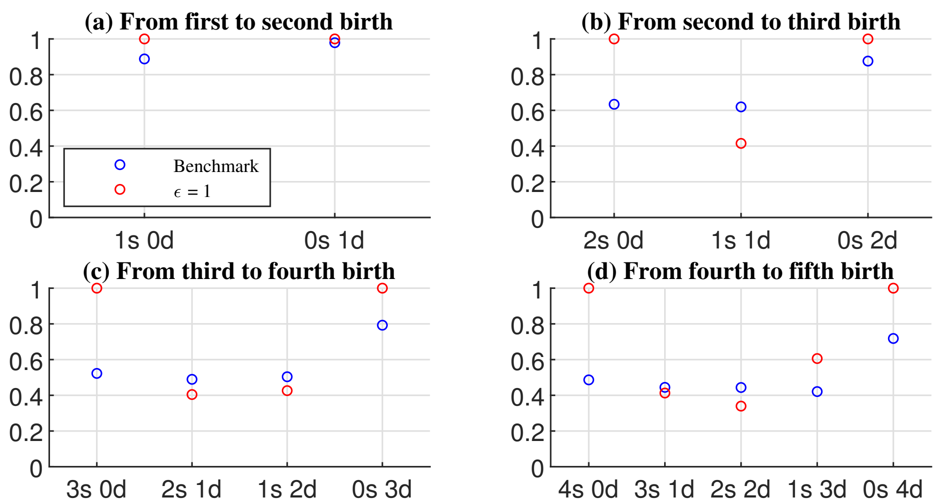



Figure A.9: Parity Progression Ratio, Conditional on the Sex Composition of Existing Children (Benchmark and Unit Substitution Elasticity).

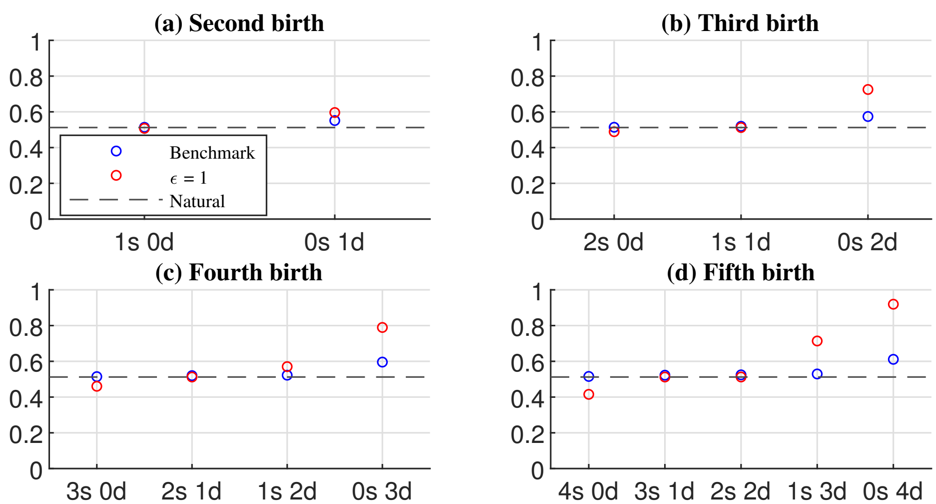



Figure A.10: Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children (Benchmark and Unit Substitution Elasticity).

## B.3 Non-targeted Moments

In this appendix, I show the model fit for some non-targeted moments, i.e., those not used in the estimation. Figure A.11 suggests that women with lower levels of education are more likely to progress to the next birth parity, while the shapes of the conditional PPR are similar across levels of education. The model does a good job of replicating these features in the data.

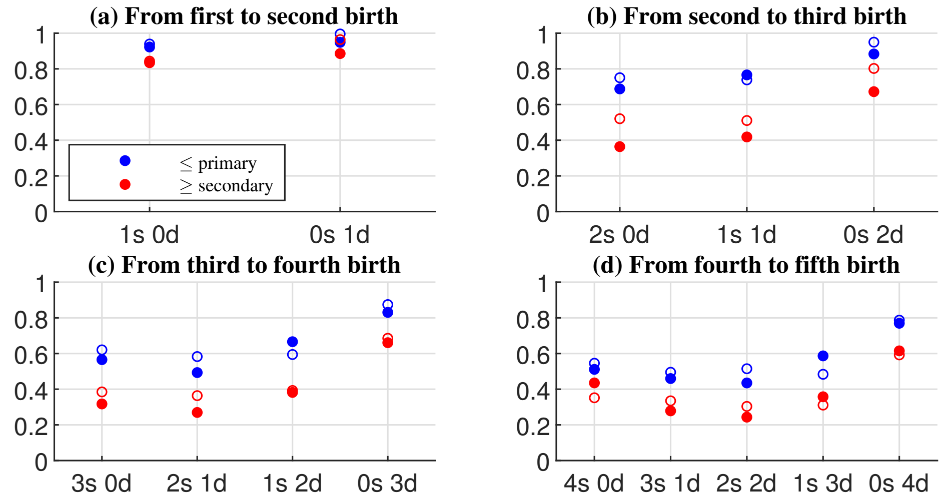



Figure A.11: Parity Progression Ratio, Conditional on the Sex Composition of Existing Children, by Women's Education (Data and Model).

Table A.9 presents the fertility rate and education expenses per son and daughter for women with different levels of education. The model effectively captures the differences by maternal education.

Table A.9: Data and Model: Selected Non-targeted Moments

| Variable | Data | Model |
|---|---|---|
| Fertility (women without secondary education, henceforth low) | 3.40 | 3.56 |
| Fertility (women with secondary education, henceforth high) | 2.46 | 2.73 |
| Education expense per son (% income, low) | 4.39 | 4.95 |
| Education expense per son (% income, high) | 6.19 | 6.88 |
| Education expense per daughter (% income, low) | 3.36 | 2.70 |
| Education expense per daughter (% income, high) | 5.74 | 4.66 |

Table A.10 reports the distribution of married men, married women, and couples by education. The model can match the data well

Table A.10: Distribution of Married Men, Women, and Couples (%)

| | | Men | |
|---|---|---|---|
| | | Primary or lower | Secondary or higher |
| Women | Primary or lower | 27.3 | 19.4 |
| | | (23.7) | (23.3) |
| | Secondary or higher | 7.0 | 46.3 |
| | | (5.6) | (47.4) |

*Notes.* Model predictions in parentheses.

# C Estimate the Model to Match Punjab Data

This appendix describes the mapping of the model to the Punjab data. Identical values are assigned to parameters whose values are chosen externally in the benchmark model. The remaining 12 parameters are estimated within the model to match the following 26 data moments in Punjab.

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

(ii) 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]

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

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

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

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

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

Note that due to the limited sample size for Punjab, 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 with precision. Therefore, they are not included in the set of targets.

The estimated parameters are presented in Table A.11. The relative weight of sons ($\omega$) is larger in Punjab than in the benchmark model, and the elasticity of substitution between sons and daughters ($\epsilon$) is smaller in Punjab, suggesting that parental decisions on fertility and sex selection are more influenced by the sex composition of existing children. As a result, the SRB is more imbalanced in Punjab.

Table A.11: Estimated Parameters for Punjab

| | Description | Benchmark | Punjab | Target |
|---|---|---|---|---|
| $\beta$ | Annual discount factor | 0.9754 | 0.9745 | Moments (vii) |
| $\nu_1$ | Weight of non-market goods when young | 0.2710 | 0.3331 | Moments (vi) |
| $\nu_2$ | Weight of non-market goods when middle-aged | 0.5782 | 0.7051 | Moments (vi) |
| $\mu_\alpha$, $\sigma_\alpha$ | Distribution of weight of children | -0.9806, 0.4998 | -0.9221, 0.3469 | Moments (i)-(v) |
| $\omega$ | Relative importance of sons | 0.5223 | 0.5468 | Moments (i)-(v) |
| $\epsilon$ | Substitution between sons and daughters | 35.24 | 24.77 | Moments (i)-(v) |
| $\mu_\xi$, $\sigma_\xi$ | Distribution of disutility from sex selection | -3.5524, 0.2330 | -3.4960, 0.1577 | Moments (i)-(v) |
| $\eta$ | Importance of education investment | 0.6093 | 0.6119 | Moments (i)-(v) |
| $e_0$ | Public education expenditure | 0.0322 | 0.0227 | Moments (i)-(v) |
| $z$ | Efficiency of human capital production | 13.83 | 14.19 | Moments (i)-(v) |

*Notes.* For Punjab, 26 moments are used to estimate 12 parameters.

Moments (iii)-(vii) in the data and model are shown in Table A.12. Moments (i) and (ii) are shown in Figures A.12 and A.13, respectively.

Table A.12: Data and Model: Selected Moments for Punjab

| Variable | Data | Model |
|---|---|---|
| Education expense per son (% income) | 6.64 | 8.23 |
| Education expense per daughter (% income) | 5.75 | 4.42 |
| Fertility difference across women's education | 0.69 | 0.66 |
| Men with secondary education (%) | 75.3 | 84.7 |
| Women with secondary education (%) | 65.5 | 63.5 |
| Female labor supply when young | 0.180 | 0.193 |
| Female labor supply when middle-aged | 0.167 | 0.165 |
| Savings rate when middle-aged (%) | 36.7 | 35.3 |

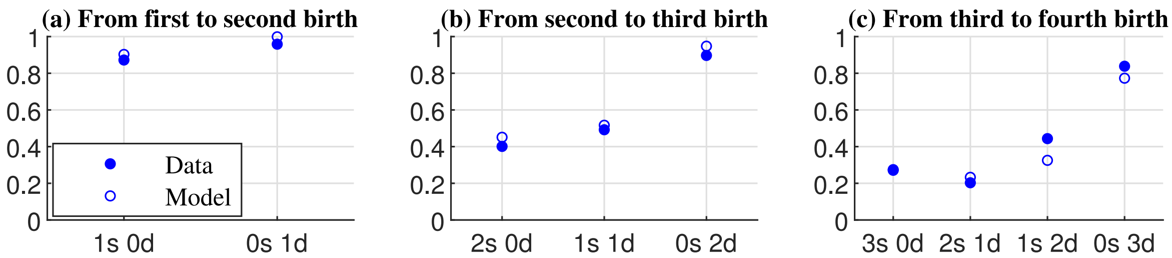



Figure A.12: Parity Progression Ratio, Conditional on the Sex Composition of Previous Children, in Punjab (Data and Model).

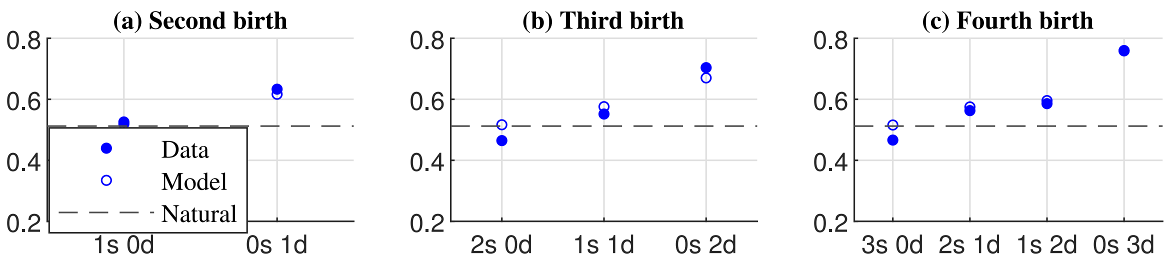



Figure A.13: Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children, in Punjab (Data and Model).

# D Additional External Validation

This appendix applies the model to simulate the impacts of a sex selection ban on the sex ratio and children's education. The model's predictions are compared to the empirical findings of Nandi and Deolalikar (2013) and Rastogi and Sharma (2022).

In 1994, the Indian government passed the Pre-Conception and Pre-Natal Diagnostic Techniques (PNDT) Act, which took effect in 1996. This act prohibits the misuse of prenatal diagnostic techniques, such as ultrasound, to determine the sex of a fetus. Violators are subject to severe penalties, including imprisonment, fines, and the suspension or cancellation of the registration of involved laboratories and clinics. Although this act applied nationwide, it was not binding in the State of Maharashtra and Jammu and Kashmir (JK), which passed their own acts in 1989 and 2002, respectively. The staggered implementation across states provides a quasi-experiment for studying its effects using a difference-in-differences approach.

Recent studies suggest that the PNDT act reduced the sex ratio while widening the gender gap in education. Nandi and Deolalikar (2013) find that the act increased female births by 1.37-2.64 per 100 male births in rural areas and 1.39 in urban areas. Rastogi and Sharma (2022) further show that the increase in female births was concentrated among families with a firstborn girl (see Table 2 in Rastogi and Sharma (2022)). In addition, they find that the act widened the gender gap in education, with girls being 2.3, 3.5, and 3.2 percentage points less likely than boys to complete Grade 10, Grade 12, and enter university, respectively (see Table 4 in Rastogi and Sharma (2022)).

In the model, I simulate the sex selection ban by increasing the monetary cost of sex selection to infinity. However, since the actual impact of the PNDT Act on the cost of sex selection is unclear, the comparison between the model and data is mainly qualitative rather than quantitative.

The results, displayed in Table A.13, suggest that a sex-selection ban can reduce the sex ratio to its natural level by design, in accordance with the findings of Nandi and Deolalikar (2013). Moreover, the increase in female births is concentrated among families with a firstborn girl, consistent with the findings of Rastogi and Sharma (2022). However, in contrast to their findings, the model suggests that the sex selection ban has minimal effects on the gender gap in education, despite a small decrease in both male and female educational attainment.[29]

Table A.13: Impacts of Sex Selection Technology

| | Benchmark | Sex selection ban (short run) | Sex selection ban (long run) |
|---|---|---|---|
| Sex ratio at birth | 1.122 | 1.050 | 1.050 |
| Probability that the second or subsequent birth is female, conditional on firstborn girl (%) | 44.0 | 48.8 | 48.8 |
| Probability that the second or subsequent birth is female, conditional on firstborn boy (%) | 48.5 | 48.8 | 48.8 |
| Men with secondary education (%) | 70.7 | 70.6 | 70.4 |
| Women with secondary education (%) | 53.0 | 53.0 | 52.7 |

*Notes.* Gender gap in secondary education is defined as the difference between the proportions of men and women with secondary education.

[29] The model also suggests that a sex selection ban has only a small effect on fertility, increasing it by 1.2%, from 3.12 to 3.16. This is in line with Li and Pantano (2023), who find that low-cost sex selection in the US can reduce fertility by 1.8%, from 2.28 to 2.24.

# E Additional Results

## E.1 Fertility and Sex Selection in Counterfactual Analyses

If economic factors are equalized between sons and daughters, the son-biased fertility stopping rule would be less evident (Figure A.14), and sex selection would not be practiced (Figure A.15).

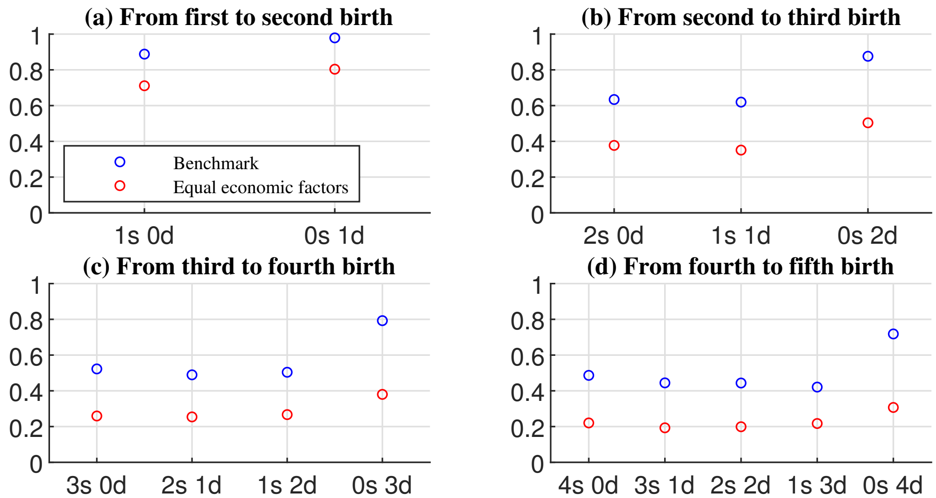



Figure A.14: Parity Progression Ratio, Conditional on the Sex Composition of Previous Children, if No Gender Difference in Economic Factors.

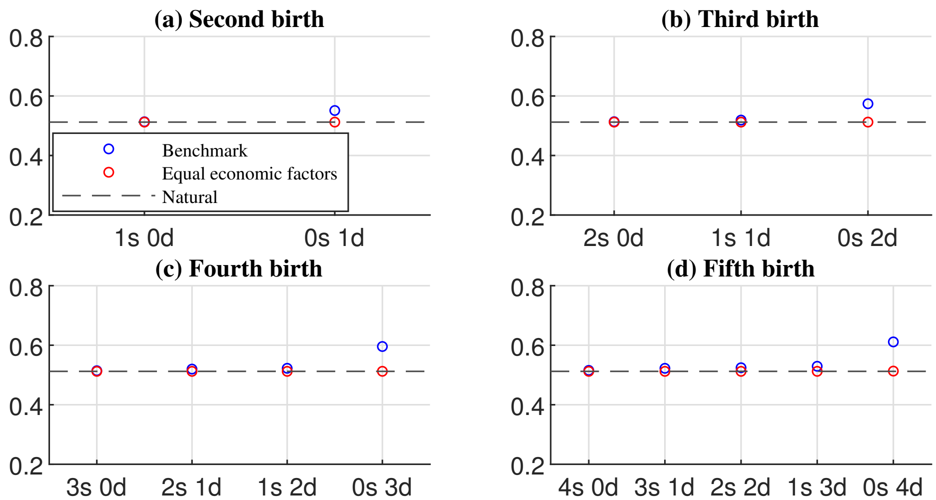



Figure A.15: Probability that the Next Birth is a Boy, Conditional on the Sex Composition of Existing Children, if No Gender Difference in Economic Factors.

## E.2 Interaction Between Old-age Support and Gender Pay Gap

Table 7 in the main text shows that the SRB will increase, rather than decrease, if there is no gender pay gap. One explanation is that the value of daughters does not increase unless their income contributes to old-age support, while the value of sons increases due to their wives' higher earnings. To provide more support for this explanation, I now remove gender disparities in old-age support and wages simultaneously. The results, displayed in Table A.14, indicate that the SRB will not increase but instead drop to 1.090, highlighting the important interaction between old-age support and gender discrimination in the labor market.

Table A.14: Joint Effects of the Gender Difference in Old-age Support and Wages

| | Benchmark | Equal support | Equal wages | Equal support and wages |
|---|---|---|---|---|
| Sex ratio at birth | 1.122 | 1.071 | 1.220 | 1.090 |
| Fertility | 3.12 | 3.02 | 2.23 | 2.16 |
| Education expense per son (% income) | 5.86 | 5.71 | 8.20 | 8.20 |
| Education expense per daughter (% income) | 3.60 | 4.41 | 5.87 | 6.64 |
| Men with secondary education (%) | 70.7 | 69.9 | 89.7 | 89.9 |
| Women with secondary education (%) | 53.0 | 60.8 | 80.2 | 84.7 |
| Female labor (young) | 0.227 | 0.247 | 0.440 | 0.451 |
| Female labor (middle-aged) | 0.262 | 0.276 | 0.449 | 0.453 |
| Household income | 100.0 | 102.3 | 120.8 | 125.8 |

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

## E.3 An Alternative Way to Equalize Marriage Payment

In the main text, to examine the effects of gender differences in marriage payments, I reduce the dowry payment by 50% but introduce an equal marriage payment for sons. This ensures gender parity in marriage costs, while leaving the pre-birth expected marriage payment (in the absence of sex selection) unchanged. Alternatively, we can eliminate dowry for daughters without introducing marriage payments for sons. The results, displayed in Table A.15, indicate that this also reduces the SRB. However, as the cost of having children has decreased, parents now have more children but invest less in their education.

Table A.15: Effects of Equalizing Marriage Payment

| | Benchmark | Both half payment | Both zero payment |
|---|---|---|---|
| Sex ratio at birth | 1.122 | 1.050 | 1.050 |
| Fertility | 3.12 | 3.16 | 4.28 |
| Education expense per son (% income) | 5.86 | 5.85 | 3.45 |
| Education expense per daughter (% income) | 3.60 | 3.57 | 1.71 |
| Men with secondary education (%) | 70.7 | 70.5 | 46.5 |
| Women with secondary education (%) | 53.0 | 52.6 | 26.9 |
| Female labor (young) | 0.227 | 0.223 | 0.153 |
| Female labor (middle-aged) | 0.262 | 0.261 | 0.238 |
| Household income | 100.0 | 100.9 | 86.3 |

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

## E.4 Alternative Subsidy Rates

In the main text, the rate of the female education subsidy is chosen to achieve the same SRB reduction as the female birth subsidy, allowing for a comparison of their financial costs. Alternatively, this rate can be chosen to equalize the labor income tax rate used to finance both policies. The results in Table A.16 show that the female education subsidy now has a smaller impact on reducing the SRB.

Table A.17 illustrates the effects of varying levels of subsidy for female births. As the subsidy increases, the SRB decreases, but fertility increases and education spending per child decreases, leading to fewer individuals completing secondary education. Additionally, the higher fertility rate reduces the labor supply of young women.

Table A.18 presents the effects of varying subsidy rates for female education. As the subsidy rate increases, the SRB decreases, education spending increases for daughters but decreases for sons, leading to the gender gap in education first narrowing and then reversing. With a higher subsidy rate, women's labor supply increases.

Table A.16: Effects of Subsidies for Female Births and Education (the Same Tax Rate)

| | Benchmark | Birth subsidy | Education Subsidy |
|---|---|---|---|
| Sex ratio at birth | 1.122 | 1.091 | 1.104 |
| Fertility | 3.12 | 3.25 | 2.95 |
| Education expense per son (% income) | 5.86 | 5.61 | 5.85 |
| Education expense per daughter (% income) | 3.60 | 3.41 | 4.77 |
| Men with secondary education (%) | 70.7 | 68.1 | 71.3 |
| Women with secondary education (%) | 53.0 | 50.1 | 63.6 |
| Female labor (young) | 0.227 | 0.213 | 0.255 |
| Female labor (middle-aged) | 0.262 | 0.265 | 0.282 |
| Household income | 100 | 98.8 | 102.9 |
| Labor income tax rate (%) | 0.00 | 0.43 | 0.43 |
| Subsidy as a share of GDP (%) | 0.00 | 0.41 | 0.40 |

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

Table A.17: Effects of the Subsidy for Female Births

| | | Subsidy value | | |
|---|---|---|---|---|
| | Benchmark | 0.004 | 0.008 | 0.012 |
| Sex ratio at birth | 1.122 | 1.115 | 1.092 | 1.076 |
| Fertility | 3.12 | 3.16 | 3.25 | 3.38 |
| Education expense per son (% income) | 5.86 | 5.79 | 5.62 | 5.37 |
| Education expense per daughter (% income) | 3.60 | 3.55 | 3.41 | 3.21 |
| Men with secondary education (%) | 70.7 | 69.9 | 68.2 | 65.8 |
| Women with secondary education (%) | 53.0 | 52.2 | 50.2 | 47.5 |
| Female labor (young) | 0.227 | 0.222 | 0.213 | 0.202 |
| Female labor (middle-aged) | 0.262 | 0.264 | 0.265 | 0.265 |
| Household income | 100.0 | 99.6 | 98.8 | 97.5 |
| Labor income tax rate (%) | 0.00 | 0.20 | 0.43 | 0.70 |
| Subsidy as a share of GDP (%) | 0.00 | 0.18 | 0.40 | 0.66 |

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

Table A.18: Effects of the Subsidy for Female Education

| | | Subsidy rate | | |
|---|---|---|---|---|
| | Benchmark | 10% | 20% | 30% |
| Sex ratio at birth | 1.122 | 1.107 | 1.076 | 1.057 |
| Fertility | 3.12 | 2.96 | 2.87 | 2.81 |
| Education expense per son (% income) | 5.86 | 5.87 | 5.61 | 5.13 |
| Education expense per daughter (% income) | 3.60 | 4.69 | 5.69 | 6.77 |
| Men with secondary education (%) | 70.7 | 71.4 | 70.0 | 66.7 |
| Women with secondary education (%) | 53.0 | 63.0 | 70.1 | 76.0 |
| Female labor (young) | 0.227 | 0.253 | 0.272 | 0.288 |
| Female labor (middle-aged) | 0.262 | 0.281 | 0.296 | 0.309 |
| Household income | 100.0 | 102.7 | 104.5 | 105.4 |
| Labor income tax rate (%) | 0.00 | 0.39 | 1.06 | 2.12 |
| Subsidy as a share of GDP (%) | 0.00 | 0.36 | 0.98 | 1.96 |

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

## E.5 Alternative Replacement Rates of Pensions

Table A.19 presents the effects of pensions at different replacement rates. As the replacement rate increases, the SRB decreases, and education investment for children declines, leading to a smaller proportion of men and women with secondary education. The effect on fertility is negative at low replacement rates, but it turns positive at high replacement rates.

Table A.19: Effects of the Pension System

| | | Replacement rate | | |
|---|---|---|---|---|
| | Benchmark | 30% | 50% | 70% |
| | (1) | (2) | (3) | (4) |
| Sex ratio at birth | 1.122 | 1.112 | 1.095 | 1.083 |
| Fertility | 3.12 | 3.10 | 3.13 | 3.15 |
| Education expense per son (% income) | 5.86 | 5.87 | 5.78 | 5.68 |
| Education expense per daughter (% income) | 3.60 | 3.53 | 3.37 | 3.26 |
| Men with secondary education (%) | 70.7 | 68.0 | 65.3 | 62.7 |
| Women with secondary education (%) | 53.0 | 49.6 | 46.2 | 43.3 |
| Female labor (young) | 0.227 | 0.250 | 0.262 | 0.274 |
| Female labor (middle-aged) | 0.262 | 0.252 | 0.242 | 0.233 |
| Saving rate when middle-aged (%) | 35.4 | 30.2 | 26.2 | 21.7 |
| Household income | 100.0 | 99.5 | 98.8 | 98.1 |
| Labor income tax rate (%) | 0.00 | 8.27 | 13.42 | 18.46 |

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

## References for Appendix


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.

Li, Q., & Pantano, J. (2023). The demographic consequences of sex-selection technology. *Quantitative Economics*, *14*(1), 309–347.

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.

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.