WHO ARE SEX SELECTING, AND WHEN?

Minu Philip*

October 2024

Click here for the latest version

Abstract

Parents with a preferential bias for sons often influence the sex composition of their children by either continuing childbearing until a desired number of sons are born, or by selectively aborting female fetuses (sex selection). I propose and empirically validate a general heuristic that characterizes such parents' son-targeting strategies in terms of *when* they decide to sex select. I show that parents typically sex select at their stated ideal number of children, consistent with a heuristic of sex selecting in order to avoid exceeding this ideal number. This pattern is further validated by exploiting the natural independence between a child's sex and the length of the preceding birth interval in the absence of sex selection: yet, among mothers at their ideal parity, a correlation emerges, consistent with the heuristic. I also provide evidence of sex selection at the first birth order, as well as of reversals in parents' decisions to sex select.

*I am grateful to Debraj Ray and Martin Rotemberg for their continuous guidance and counsel, which have greatly refined this project to its current form. This project has also benefited from conversations with Alberto Bisin, Katarína Borovičková, Guillaume Fréchette, Manisha Shah, and Prashant Bharadwaj.

I. INTRODUCTION

Son-preference refers to a widespread preferential bias for male children. Motivated by such preferences, parents are known to influence the sex-composition of their children—a phenomenon known as "son-targeting." This involves the use of one or both methods: *differential stopping*, referring to continued childbearing until a desired number of sons are born; and *sex selection*, the selective abortion of female fetuses in-utero. These practices distort birth sex ratios, fertility rates, and the order and spacing of children; with far-reaching consequences for population demographics, marriage market dynamics, maternal health, and child well-being. While the literature extensively documents the use and consequences of the two son-targeting methods, less is known about the timing and conditions under which parents switch to sex selection rather than relying on the natural odds of male births.

This paper identifies a general heuristic that characterizes parental son-targeting strategies, specifically indicating when son-preferring parents choose to engage in sex selection. Using retrospective birth histories of Indian women from three rounds of the National Family Health Surveys (NFHS),[1] I analyze the son-targeting methods that parents employ and the specific decision rules they follow. India provides a particularly suitable context for this study due to the widespread prevalence of son-preference, reflected in its skewed sex ratios, combined with an ongoing fertility transition that generates sufficient variation in fertility rates to support detailed analysis—unlike many countries with either very high or low fertility.

When making fertility decisions under son-preference, parents face a choice: they can either rely on the natural probability of conceiving a child of a desired sex—risking an unwanted birth—or incur the costs of sex selection to avoid unwanted births and potentially try again. Since parental interventions aimed at influencing sex composition are neither directly observed nor self-reported, identifying a heuristic that governs such behavior is inherently challenging.[2] Any attempt to uncover this heuristic must therefore necessarily rely on inferring the use of these methods from realized birth histories. However, even with detailed birth histories, distinguishing instances of sex selection from differential fertility stopping is not straightforward, as both methods are driven by parents' unmet desire for sons and can generate similar birth history patterns. Consequently, the existing literature has predominantly studied these two methods separately.[3]

While both methods are driven by a desire for more sons, sex selection is specifically necessitated by a high cost of continued childbearing. Since this cost typically increases with each subsequent birth, parents who sex select will not revert to differential stopping thereafter, making any switch between these methods effectively one-way (Baland et al., 2023). I leverage this distinction in parental motivation to identify when parents switch to sex selection. More precisely, I define the *relative birth order* for each child, representing the difference between the child's birth order and the mother's self-reported ideal number of children (ideal

[1] NFHS is the India-specific implementation of the Demographic and Health Survey (DHS) framework.

[2] Parents typically do not acknowledge use of sex selection, as prenatal sex determination is prohibited in India under the Pre-Conception and Prenatal Diagnostic Techniques (PCPNDT) Act of 1994.

[3] An exception is Baland et al. (2023), which develops a unified framework to study both methods and proposes tests to detect the use of differential stopping and sex selection at various birth orders.

parity). For instance, a mother with an ideal parity of two who has three children, born at orders 1, 2, and 3, would have relative birth orders of -1, 0, and 1, respectively; with the last birth being *one more than* what the mother had desired. Relative birth orders reflect the increasing cost associated with each successive birth, enabling the identification of when these costs become large enough to prompt sex selection.

Comparing the proportion of sons born at various relative birth orders, I find a significantly high proportion of male births when mothers are at their ideal parity. I confirm that this pattern is not driven by parents simply reporting their actual parity as their ex-post rationalized ideal number. Because a parent's ideal parity should not naturally affect the probability of having a son at any birth order, this pattern strongly suggests that sex selection occurs at the ideal parity. This supports a heuristic by which son-preferring parents engage in sex selection when approaching their ideal parity, to avoid exceeding it. The same pattern holds consistently across rural and urban areas, and over three rounds—NFHS-3 (2005–06), NFHS-4 (2015–16), and NFHS-5 (2019–21)—covering cohorts of women who completed childbearing in different decades. Minor yet noteworthy differences emerge by education and wealth: more educated and wealthier mothers are less likely to exceed their ideal parity compared with mothers who have only primary or no education, and hence display a stronger tendency to sex-select. Heterogeneity is also observed along geographical regions in India, synonymous with cultural divisions and the prevalence of son-preference.

To empirically validate this heuristic identified from patterns of sex ratios across relative birth orders, I exploit the human sex-determination system that renders a newborn's sex to be naturally independent of the sex composition of previously born children. In India, where the revelation of a fetus's sex in-utero is illegal, sex should be revealed *legally* only at the time of birth, i.e. at *end* of the *preceding* inter-birth interval. The length of the preceding inter-birth interval and the sex of newborn born at the end of it, should therefore be uncorrelated and independent of the composition of children already born to the mother.

Sex selection, however, artificially increases the observed birth interval preceding a live male birth. This increase arises because fetal sex cannot be determined until at least 10–12 weeks of gestation, and additional time is required for post-abortion recovery and the subsequent conception attempt (Pörtner, 2016, 2022). Alternatively, if parents anticipate these delays and respond either by initiating conception earlier when planning sex selection or by reversing sex-selection decisions after unsuccessful attempts—girls may be born following longer intervals. Any correlation between the sex of a child and the length of the preceding interval therefore indicates parental manipulation. Even if other factors could generate such a correlation, they should affect all birth histories equally. Consequently, any difference in the correlation between a child's sex and the length of the preceding birth interval *across* birth histories must be attributed to parental manipulation—since only the incentives to sex select vary with birth history, while the effects of any confounding factors would not.

Exploiting the natural independence between a child's sex and the preceding birth interval, I identify the use of sex selection by examining whether the sex of a newborn "affects" the length of the interval preceding its birth. For this, I use a difference-in difference specification that compares the length of the inter-birth interval preceding male and female births across different birth histories. The identifying assumption here—akin to the parallel trends assumption—is that in the absence of sex selection, the difference in birth-intervals preceding male and female births shouldn't be different across birth histories. This "assumption", to our convenience, is guaranteed by human biology as sex is determined independently of previous births. I then use this difference-in-difference specification to test for sex selection among sub-samples of parents that the heuristic suggests would engage in sex selection versus those who, according to the heuristic, would not. I find that at any birth order, sex selection is driven by the mothers who report an ideal parity equal to the birth order, consistent with the heuristic. Extending the analysis to the first birth order, I also find evidence of mothers with an ideal parity of one resorting to sex selection as early as at the first birth order.

This research contributes to a vast literature on son-preference and associated son-targeting behaviors, briefly reviewed in Section II. In contrast to most studies in this scholarship which analyze either differential stopping or sex selection in isolation, I examine parents' son-targeting strategies allowing for the possibility that both methods may be employed in combination. This approach is closely related to Baland et al. (2023), who also develop a unified framework to study the use of both methods and propose empirical measures to detect their use. In their formalization of fertility decision-making, Baland et al. (2023) characterize the timing of sex selection relative to a theorized "maximum number of pregnancies in a family." The findings in this paper suggest a sharper behavioral interpretation: most parents appear to treat their ideal number of children as a binding maximum and face substantial costs of deviating from it, actively intervening to align fertility outcomes with this target.

Another important departure from the existing literature is the documentation of sex selection at the first birth order. Prior studies have relied on the assumption that the sex of the firstborn is as good as random, on the premise that parents do not engage in sex selection at the first birth order. The findings in this paper, however, indicate that parents whose ideal family size is one may indeed be likely to sex select at the first birth order. Moreover, the analysis of preceding birth intervals—an approach also employed in Pörtner (2022)—sheds light on the dynamics of fertility decisions between two live births, revealing how some parents may even reconsider or reverse earlier decisions to sex select when the spacing between births becomes "too large."

By identifying a general heuristic that guides son-targeting strategies and demonstrating its close link to ideal parity, this paper cautions against any simplistic expectations that sex-selective practices will diminish with economic development or rising education. Given the negative correlation between desired parity and both economic development and maternal education, the challenge of sex selection is likely to become more salient rather than less. Policy design must therefore account for parents' parity preferences when designing interventions that either incentivize or penalize certain fertility choices.

The remainder of this paper is organized as follows. Section II provides reviews the two son-targeting methods, describes the data used in the analysis, and examines the prevalence of the two methods. Section III outlines the conceptual framework and empirical strategy, which are then applied in Section IV to characterize parents' son-targeting behavior and to identify and validate the heuristic of sex selection. Finally, Section V concludes.

II. BACKGROUND AND DATA

II.A. Son-targeting methods and their indications

A preferential bias for sons over daughters is globally widespread, and stems generally from a combination of cultural, religious and economic factors. In the case of India, patrilocal kinship and the custom of dowry payments precipitate sentiments of girl children being non-members at their natal homes ("*paraya dhan*") and a liability until they're married and sent away. The prevalent gendered division of labor reinforces this bias as sons are perceived as the primary providers of financial support to parents in old age. Religious rites and inheritance laws also grant sons important roles, leaving parents wanting a son to carry forward their family name, lineage, and wealth. The geographical prevalence of son preference in India, therefore, mirrors the geographical divide in religious and cultural practices. Son-preference is more widespread in the northern and western states of India compared to the historically matrilineal southern and north-eastern states. Pande & Astone (2007) and Jayachandran (2015) highlight the importance of a range of socioeconomic and sociocultural factors in shaping son preference, operating at both individual and community levels.

Motivated by son-preference, parents influence the sex composition of their children through differential fertility stopping and sex selection. The use of these son-targeting methods has been extensively documented in the literature, with each method identified through distinct empirical patterns, as outlined below.

Differential stopping is indicated by high parity progression ratios—that is, a greater likelihood of parents continuing to have children following birth histories with fewer sons compared to other histories at the same parity (Das, 1987). It results in girl children having more siblings on average, and being the older sibling in these large families (Basu & Jong, 2010). Baland et al. (2020) formalize these patterns to identify countries where stopping rules are prevalent; such countries display a consistent trend where, at any given birth order, girls are more likely than boys to have younger siblings. Differential stopping also influences birth spacing: parents with an unmet desire for sons often attempt conception sooner, leading to shorter intervals following female births (Javed & Mughal, 2020). This, in turn, shortens breastfeeding durations for girls born at lower birth orders (Jayachandran & Kuziemko, 2011). Consequently, girl children in families practicing differential stopping face higher sibling rivalry and competition for nutritional investments (Jayachandran & Pande, 2017).

sex selection, on the other hand, is indicated by skewed sex ratios (Sen, 1990, 1992; Jha et al., 2006). It also manifests in family-level patterns, such as a higher proportion of girls among the older siblings of boys compared to those of girls (Baland et al., 2023). Another empirical marker is the longer inter-birth intervals preceding male births relative to female births (Pörtner, 2016, 2022). This pattern arises because fetal sex can be determined only after around 12 weeks of pregnancy, and when parents choose to abort and try again for a son, the combined duration until next conception and birth results in a longer inter-birth interval before the next (male) birth.

While the literature has examined differential stopping and sex selection extensively, it has largely studied the two methods in isolation. There is limited evidence on whether parents consistently adopt one method over the other, or whether they alternate between the two based on changing circumstances. In particular, the literature has been mostly silent on whether sex selection is used exclusively or in conjunction with differential stopping, and—crucially—on the question of timing: at what point in their fertility histories do parents choose to resort to sex selection, if at all?

II.B. Data

To study parents' choices between the two son-targeting methods and how these relate to their fertility preferences and birth outcomes, I analyze retrospective birth histories of Indian women aged 15–49, as recorded in the National Family Health Survey (NFHS). The NFHS collects detailed information on women's birth histories, fertility intentions, and reproductive health practices. I use data from three survey rounds: NFHS-3 (2005–06), NFHS-4 (2015–16), and NFHS-5 (2019–21). For each round, I restrict the sample to women who have had at least one birth and report having completed childbearing.[4] I then exclude women who had multiple births in a single pregnancy. Each survey round thus provides a cross-section of women who completed childbearing in different decades. The NFHS-3 sample includes 67,451 women whose childbearing occurred primarily between 1978 and 2005; NFHS-4 includes 366,224 women with births between 1987 and 2015; and NFHS-5 covers 375,004 women with births between 1991 and 2019.[5]

Respondents provide detailed information for each birth, including the sex of the child, birth order, and the length of the preceding and succeeding inter-birth interval.[6] The surveys also collect a wide range of individual and household-level characteristics, such as the respondent's age, level of education, wealth index,[7] contraceptive use, and reproductive health practices. Importantly, the NFHS also gathers information on fertility preferences. Respondents are asked to report their ideal parity and ideal composition of children—that is, the number of sons, daughters, and children of either sex that *"the respondent would like to have in her whole life, irrespective of the number she already has."*

Table 1 summarizes the fertility preferences and actual birth outcomes of Indian women over the three rounds. While women's age at the beginning and end of childbearing has largely remained the same over the years, there has been in steep decline in actual parity. This decline cannot be attributed to a commensurate change in women's preferred or ideal parity, which has instead remained fairly stable. It must then be that either parents are able to achieve their desired composition at earlier parities, or have increasingly become unwilling to trade-off low parity for a more desired sex composition.

[4] Women are asked whether they would like to have another child. Those who respond that they do not want more children or are unable to have more—due to sterilization or infecundity—are classified as having completed fertility.

[5] Birth year ranges are based on the 5th and 95th percentiles of birth years reported.

[6] For all births reported, the survey confirms whether the child is alive at the time of the survey and if not, the age at death. In case the respondent fails to report births that were subject to infanticide, it would be difficult to separate instances of unreported female infanticide from longer birth intervals.

[7] The wealth index is a composite measure of the household's living standard based on aspects such as the household's ownership of assets, access to living amenities, and type of housing, among others. It is defined on a continuous scale and is used in the data to categorize households into quintiles of relative wealth.

Table 1 :
Descriptive Statistics

	NFHS Rounds		
	3 2005-06	4 2015-16	5 2019-21
Median age at first birth	20	20	20
Median age at last birth	26	26	25
Total Parity (Mean)	3.278	3.016	2.796
Proportion of sons (Mean)	0.551	0.565	0.563
Ideal parity (Mean)	2.399	2.399	2.322
Ideal number of sons (Mean)	1.136	1.225	1.150
Ever used modern contraception (excludes permanent methods)	0.393	0.242	0.348
No. of respondents	67,451	366,224	375,004

Notes: This table summarizes the fertility preferences and actual birth outcomes of women who report having completed childbearing. The data is drawn from the 2005-06, 2015-16, and 2019-21 rounds of the National Family Health Survey for India.

An important concern that may arise here is over the reliability and interpretation of reported ideal parity, as these responses may reflect an ex-post rationalization based on a woman's actual number of children. To address this, I begin by examining the correlation between actual and reported ideal parity, finding it to be moderate at 0.51. This suggests that reported ideals are not simply mirroring realized outcomes. Moreover, the correlation is even lower when there's multiple births at a pregnancy, further indicating that stated 'ideal' parity does not closely track actual fertility (see Appendix Table B1).[8]

I interpret the reported ideal parity figures as reflecting the fertility preferences that respondents in my sample used to guide their fertility choices. This interpretation, however, requires some clarification. First, respondents' preferences for ideal family size and sex composition may have evolved over time as they experienced successive births. Since my sample is restricted to women who have completed childbearing, I take their reported preferences as those relevant to decisions made around their last birth—specifically, the decision to stop having children. Second, fertility decisions are rarely shaped by the preferences of the mother alone; they are often also influenced by her partner and other family members. Given the sample restriction to women who have completed their fertility, it is reasonable to assume that they are aware of their partner's preferences. More than 75% of women in the NFHS-3

[8] Jayachandran & Kuziemko (2011) also use reported ideal parity in their analyses and discuss similar concerns regarding its interpretation.

sample report that their partner "wants the same number of children [...] that she wants herself," and this statistic rises to about 85% in the NFHS-4 and NFHS-5 samples. Whether this alignment reflects pre-existing agreement or results from SOME intra-household bargaining is less important in my context since the focus is on capturing preferences that ultimately drive fertility decisions.

Section III explains how I use data on birth outcomes and fertility preferences to analyze parental choices between the two son-targeting methods and identify a heuristic. Before undertaking that analysis, however, it is important to first establish the continued relevance of such an investigation by examining the current prevalence of these song-targeting practices. This next section analyzes the extent to which differential stopping and sex selection remain in use and whether their prevalence has changed over time.

II.C. Prevalence of son-targeting methods: Crude estimates

Although the use of differential stopping or sex selection is not directly reported or admitted to, it is possible, under certain assumptions, to deduce crude proportions of parents employing each method. For this exercise, I follow the following definitions of the two son-targeting methods:

Let h represent a birth history (the sex-sequence of children born to a parent),

Definition 1 *: **Differential stopping behavior***
If there exists a pair of birth histories $h \neq h'$, each with the same number of children (k) but differing in sex-composition, such that the likelihood of stopping childbearing after h differs from that after h', this is indicative of differential stopping behavior at birth order k.

Definition 2 *: **Sex selection***
If there exists any positive measure of parents selectively eliminating female births conditional on birth history h, it is indicative of sex selection following history h.

Suppose the natural probability of giving birth to a male child is known and identical across all mothers, denoted by π. Further, assume that mothers who engage in sex selection are willing to selectively abort female fetuses up to T times, if necessary, in order to conceive a male child. If a male birth does not occur after T such attempts, they abandon further attempts and stop conceiving altogether.

If N_h denoted the number of mothers whose birth history begins with the sequence h,
μ_h denoted the proportion of mothers who stop childbearing following history h, and
λ_h the proportion of mothers who sex selectively abort following history h;

then following history h,

$$\begin{aligned} \pi(1-\mu_h)N_h &\text{ mothers would birth a successive male child naturally} \\ (1-\pi)(1-\lambda_h)(1-\mu_h)N_h &\text{ mothers would birth a successive female child} \\ \lambda_h(1-\pi)(1-\mu_h)N_h &\text{ mothers would sex select} \end{aligned}$$

Of the $\lambda_h(1-\pi)(1-\mu_h)N_h$ mothers who would selectively abort, a proportion π would birth male children in their second attempt, followed by a proportion $\pi(1-\pi)$ who would birth a male child in their third attempt and so on, up to $T < \infty$ attempts.

With no take-back or reversal of the decision to sex select, total male births at the successive order will be

$$\underbrace{\pi(1-\mu_h)N_h}_{\text{male birth at first attempt}} + \underbrace{\lambda_h(1-\pi)(1-\mu_h)N_h\left[\pi + \pi(1-\pi) + \pi(1-\pi)^2 + ... + \pi(1-\pi)^{(T-1)}\right]}_{\text{male births after selective abortions}}$$

$$= \left\{\pi + \lambda_h(1-\pi)[1-(1-\pi)^T]\right\}(1-\mu_h)N_h$$

Total successive births will be

$$\left[\underbrace{(1-\pi)(1-\lambda_h)}_{\text{female}} + \underbrace{\pi + \lambda_h(1-\pi)[1-(1-\pi)^T]}_{\text{male}}\right](1-\mu_h)N_h$$

The proportion of female successive births is hence given by $\frac{\left[(1-\pi)(1-\lambda_h)\right]}{\left[1-\lambda_h(1-\pi)^{T+1}\right]}$.

Let $\hat{p_h}$ be the proportion of male successive births observed in the sample following history h. Equating $1-\hat{p_h}$ to the theoretical proportion of female successive births, we can obtain a crude *estimator for the proportion of sex selectors* at history h, given by

$$\hat{\lambda_h} = \frac{(1-\pi)-(1-\hat{p_h})}{(1-\pi)-(1-\hat{p_h})(1-\pi)^{T+1}} = \frac{\hat{p_h}-\pi}{(1-\pi)-(1-\hat{p_h})(1-\pi)^{T+1}} \tag{1}$$

It is the ratio of excess male births scaled by the probability of male birth within T sex selection attempts by parents whose first conception after history h is of a female fetus. The higher the number of attempts parents are willing to make, the lower is the proportion of sex selectors given excess male births.

With the proportion of sex selectors known, we can also back out the proportion of parents who decide to stop childbearing following history h—the fertility stoppers. Had all non-sex-selecting parents following history h conceived again, there would have been $(1-\pi)(1-\lambda_h)N_h$ successive female births. The crude proportion of fertility stoppers after h can therefore be backed out by the percentage shortfall in the number of actual successive girls born ($N_{h,\mathtt{F}}$) from the number of girls who *would have been born* had all non-sex-selecting parents had another child.

$$\hat{\mu_h} = 1 - \frac{N_{h,\mathtt{F}}}{(1-\pi)N_h(1-\hat{\lambda_h})} \tag{2}$$

Table 2 reports crude estimates of the proportion of fertility stoppers and sex selectors based on data from the three rounds. Positive proportions of sex selectors is indicative of use of sex selection; and *differences* in the proportion of fertility stoppers at the same parity but for histories with different sex-composition, is indicative of differential stopping behavior. Since live births are known to consistently be male-biased in nature, $\pi = 0.512$ is used to generate the estimates (Orzack et al., 2015). Table 2 assumes the maximum number of selection attempts to be $T = 3$. Using a smaller value of T would increase the estimated proportion of sex selecting parents, while allowing more than three attempts would not significantly reduce the estimates. It is important to acknowledge that there is no consensus on what the appropriate values of π and T should be, or on whether these parameters should be uniform across all parents. Nevertheless, this assumption is made here solely for the purpose of exploratory comparison. Although proportions are ideally supposed to be between 0 and 1, the estimates in Table 2—based on crude calculations—sometimes take negative values when they are close to zero. The objective here is not to take the estimates at face value, but to make relative comparisons across histories and across time periods to deduce ballpark proportions of parents using the two son-targeting methods.

TABLE 2 :
Crude estimates of sex selectors and fertility stoppers

	Proportion of *Stoppers*	Proportion of *sex selectors*	Proportion of *Stoppers*	Proportion of *sex selectors*	Proportion of *Stoppers*	Proportion of *sex selectors*
NFHS Round	2004-05		2015-16		2019-21	
-	N/A	0.04	N/A	0.07	N/A	0.06
`M`	0.10	0.01	0.11	0.01	0.12	-0.01
`F`	0.07	0.10	0.06	0.11	0.07	0.11
`MM`	0.39	0.00	0.48	-0.02	0.54	-0.05
`MF`	0.36	0.04	0.42	0.04	0.49	0.02
`FM`	0.37	0.04	0.44	0.05	0.51	0.05
`FF`	0.19	0.12	0.19	0.15	0.23	0.17
`MMM`	0.43	-0.02	0.50	-0.02	0.57	-0.05
`MMF`	0.47	0.00	0.55	-0.02	0.62	-0.02
`MFM`	0.50	0.02	0.59	0.00	0.66	-0.04
`FMM`	0.50	-0.01	0.60	0.00	0.67	-0.02
`FFM`	0.44	0.03	0.50	0.05	0.58	0.09
`FMF`	0.35	0.06	0.40	0.05	0.46	0.07
`MFF`	0.34	0.08	0.39	0.06	0.46	0.06
`FFF`	0.21	0.11	0.21	0.15	0.25	0.15
Observations	67451		366224		375004	

Notes: This table reports crude estimates of the proportion of sex selectors and fertility stoppers following select birth histories, using Equations (1) and (2). In the calculation of these numbers, I assume: π, the natural probability of conceiving male child $= 0.512$, and T, the maximum number of abortions that sex selectors are willing to undergo $= 3$. Since these estimates are based on crude estimators, select values close to 0 are negative.

Table 2 suggests that for all histories, the proportion of fertility stoppers has increased over time, as can be reconciled with the decline in total parity over time noted earlier. This increase in the proportion of fertility stoppers is relatively modest for histories with only female births i.e. `F, FF, FF`—indicating the continued prevalence of differential stopping.

The prevalence of sex selection, on the other hand, is neither inconsequential nor has it shown a meaningful decline. Consider the birth history `F` (i.e., a firstborn female). Roughly 93% of parents with this history proceed to have a second child, and among them, about 11% appear to have resorted to sex selection to ensure the birth of a male child. Importantly, this 11% reflects only those who conceived a female fetus and then acted upon their preference. The true proportion of parents who would have been willing to sex select before knowing the sex of the fetus is likely much higher. The estimates of sex-selectors reported in Table 2 are also broadly consistent with previous findings: Bhalotra & Cochrane (2010) estimate 11.6% of second-order births and 15.3% of third-order births to have been subject to prenatal sex selection between 1995 and 2005.

Contrary to claims made previously in the literature, the estimates also suggest that parents may be sex selecting even at the first birth order; I confirm this observation more rigorously in Section IV.B.

III. EMPIRICAL STRATEGY

Consider parents whose fertility preferences are characterized by two parameters: n^*, the ideal number of children (ideal parity), and b^*, the ideal number of boys.[9] At each birth order, parents choose whether to stop childbearing, continue childbearing while relying on natural odds of birthing either sex, or continue while using sex selection to avoid an undesired birth. Up to the first $n^* - b^*$ children, parents would proceed without intervention relying on natural odds. Thereafter, their behavior depends on whether there is an unmet desire for sons—i.e., if fewer than b^* boys have been born so far. Let $\hat{b}$ be the deficit of sons from the desired number at this time. If there is no unmet desire ($\hat{b} = 0$), parents would continue until n^* and stop. If $\hat{b} > 0$, they may either proceed without intervention, accepting the risk of an undesired sex, or use sex selection to avoid it. As Baland et al. (2023) note, both strategies involve having another child but differ in execution: the former relies on natural odds, while the latter uses abortion to prevent the birth of an undesired sex.

Theoretically, if the cost of sex selection were zero, the earliest birth order at which a parent would sex-select is $n^* - \hat{b} + 1$. In reality, the cost of sex selection—whether monetary or moral—is rarely zero, and parents will choose it only when this cost is lower than that of having a successive child of the undesired sex. The decision to sex-select is thus necessitated by a high cost of additional births. Since this cost typically increases with each subsequent birth, parents who sex-select will not revert to differential stopping thereafter. If anything, parents may switch from differential stopping to sex selection at later orders, but not the reverse.[10]

Parents' son-targeting strategy can therefore be summarized as a general heuristic describing *when* a parent decides to sex-select: parents have $n^* - m_1$ children naturally and use sex

[9] Appendix C outlines a conceptual framework for parental fertility preferences and provides an interpretation for these 'ideal' numbers.

[10] The implicit assumption is that the cost of sex selection is non-increasing with birth order, or, if it does increase, it does so at a slower rate than the cost of an additional birth.

selection for up to m_2 additional births, resulting in a total of $n^* - m_1 + m_2$ children. Since total fertility would always be $\geq n^*$, it must be that $m_2 \geq m_1$. If $m_1 < 0$ and $m_2 = 0$, the parent relies exclusively on differential stopping. If $m_1 \geq 0$ and $m_2 > 0$, the parent relies exclusively on sex selection. Since the timing of sex selection depends on the birth order at which the relative cost of continuing naturally exceeds that of sex selection—a trade-off that reflects underlying fertility preferences—I describe the decision-making heuristic in terms of how far a birth is from the mother's ideal parity. To formalize this, I define relative birth orders as below.

Definition 3 ***Relative birth order with respect to ideal parity***
Let k denote the birth order of a child, and n^* denote the ideal parity of the parent; then the relative birth order of the child with respect to the mother's ideal parity is defined as $j^* \equiv k - n^*$.

Relative birth order reflects the rising cost of relying on natural odds. For example, for a parent whose ideal number of children is $n^* = 2$ but who has actually given birth to three children, her youngest child is born at $j^* = 1$ — which is *one over the mother's ideal number*. As j^* increases, the cost of an additional birth increases. For $j^* \leq 0$, there is net benefit from each additional birth; but $j^* \geq 1$ onward, each additional birth has a positive cost potentially increasing in j^*. Parents would give birth at $j^* > 0$ only if the expected gains from a more desirable composition were high enough to offset the cost of the additional birth.

A key advantage of defining birth order relative to ideal parity is that it enables meaningful comparisons of births occurring at different birth orders across mothers with varying fertility preferences. For example, a child born at the second birth order has a different significance for a mother whose ideal parity is two ($k = 2; j^* = 0$) than for a mother whose ideal parity is four ($k = 2; j^* = -2$). At the same time, this child can be compared—regarding parental fertility incentives—to a child born at the fourth birth order to a mother with an ideal parity of four ($k = 4; j^* = 0$), highlighting how relative birth order captures the strategic context of fertility decisions.

I characterize the heuristic of sex selection by the relative birth orders at which parents choose to sex-select. For a heuristic in which parents have $n^* - m_1$ children naturally and use sex selection for up to m_2 additional births, sex-selection would occur at relative birth orders $j^* = -m_1 + 1, \ldots, -m_1 + m_2$.

III.A. Identifying a Heuristic of Sex-Selection

To identify which relative birth orders parents sex-select at, I examine sex-ratios at relative birth orders. The key intuition is that the biological realization of a child's sex is independent of parental preferences (n^*), and thus independent of the relative birth order j^*. In the absence of sex selection, male-to-female birth ratios should equal the natural sex ratio across all j^* values. Any systematic deviations from this pattern can therefore be attributed unambiguously to sex selection.

Proposition: *The use of sex-selection can be unambiguously inferred from skewed sex-ratios at relative birth orders (j^*).*

1. *In the absence of son-targeting, parents would birth up to their ideal parity ($N = n^*$), and the proportion of sons at all relative orders $j^* \leq 0$ would be equal to the natural probability of birthing a male child.*

2. *Differential stopping does not alter the proportion of sons born at relative orders (j^*). Only when sex selection is used at a particular j^* does the proportion of male births at that order increase.*

The first statement in the proposition is straightforward; let us focus on the second. Consider parents who only practice differential stopping. For simplicity, assume that they continue childbearing until exactly b^* boys are secured, no matter the number of births it takes. These parents could produce high male-to-female birth ratios at the last birth because they stop childbearing only once a son (or the desired number of sons) is born. Parents who couldn't naturally birth b^* sons in their first n^* attempts, would necessarily birth $N > n^*$ children. For them, their youngest child would always be a boy—their b^{*th} son.

Suppose $n^* = 1, b^* = 1$. Parents who birth more than $n^* = 1$ children would only be those who did not have a boy on their first attempt. With $b^* = 1$, all children except the youngest child must necessarily be girls. For $n^* = 2, b^* = 1$, parents with a male firstborn could have their youngest child born at k = 2 be a daughter. But those have three or more children must necessarily have given birth to a daughter first. More generally, the proportion of sons at lower birth orders would decrease as the family size increases, since we are selectively observing parents who chose to continue childbearing precisely because they had not yet reached their desired number of sons. This selection into higher parity skews sex ratios *when disaggregated by actual parity.* If we attempted to account for heterogeneity in parents' fertility preferences by disaggregating the sample by actual parity, then differential stopping alone could also skew sex-ratios, potentially leading to mistaken inferences about the presence of sex selection; as in Figure 1, Panel (B). By contrast, analyzing sex ratios at relative birth orders avoids this problem. As long as sex-selection is not resorted to, the proportion of sons at all relative orders j^* would still be equal to the natural probability of giving birth to a boy. This is because sex is determined naturally and independently at each birth order no matter the value of n^* or k; see Figure 1 Panel (A).

Since parents typically care about both parity and composition, the relative weight they place on each determines the heuristic they follow—tolerating some deviation from n^* to be closer to b^*, or vice versa. I describe a heuristic in which parents have $n^* - m_1$ children naturally and rely on sex selection for up to m_2 additional births, resulting in a total of $n^* - m_1 + m_2$ children. This would correspond to observing skewed sex ratios at relative birth orders $j^* = -m_1 + 1, \ldots, -m_1 + m_2$. Figure 2 illustrates this in two cases: one where parents sex-select only once at birth order $n^* + 1$ (i.e., $m_1 = 0$, $m_2 = 1$; and $j^* = 1$), and another where they sex-select only once at n^* (i.e., $m_1 = 1$, $m_2 = 1$; and $j^* = 0$). In the absence of sex-selection, the proportion of sons at any relative order j^* will equal the natural probability of conceiving a boy. A spike in the proportion of sons at a given relative birth order j^* can only occur if parents sex-select at that order.

There are, however, two key challenges to the approach of identifying the heuristic of sex selection as described above. First, there is no consensus on the natural probability of a

FIGURE 1 :

Sex-ratios with only differential stopping ($m_1 < 0, m_2 = 0$)

(A) **At relative birth orders** (B) **Comparison Exercise**

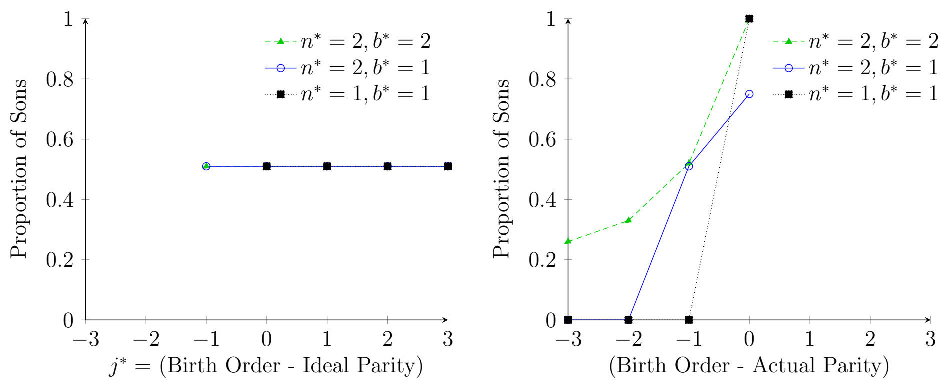
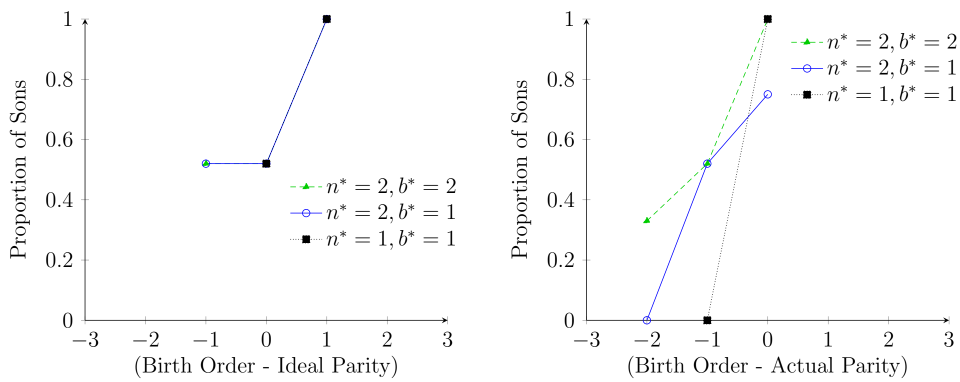


FIGURE 2 :

Sex-ratios with sex-selection

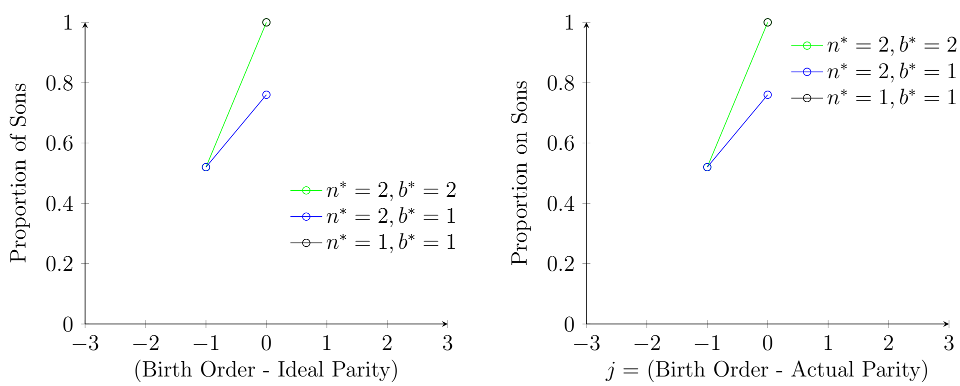
(A)

$m_1 = 0, m_2 = 1$: **Sex-select once at** $n^* + 1$ **if no. of boys** $< b^*$



(B)

$m_1 = 1, m_2 = 1$: **Sex-select once at** n^* **if no. of boys** $< b^*$



male birth to benchmark sex-ratios against. The natural sex ratio at birth is typically skewed slightly in favor of males, largely due to higher female fetal mortality during gestation (Orzack et al., 2015). Moreover, this ratio has been shown to vary by maternal age, geography, ethnicity, socio-economic status, and environmental conditions (Anderson & Ray, 2010; Chao et al., 2019); all of which also likely correlate with son-preference and the tendency to sex select. Therefore, it is not sufficient to infer sex selection solely based on observing a 'high' proportion of sons at any relative birth order j^*.

Second, the approach relies on the assumption that reported ideal parity represents a genuine prior preference, not merely a retrospective justification of realized fertility. If indeed reported ideal parity was just an ex-post rationalized reportage of actual parity, the exercise would produce the pattern of the proportion of sons by relative birth order as shown in panel B of the figures above. While I conduct robustness checks to assess the plausibility of the assumption that reported ideal parity reflects more than parents simply quoting their actual parity, a more rigorous validation requires an independent indicator.

To this end, I use an additional observed statistic to validate the heuristic: the length of the inter-birth interval, as is described next.

III.B. Validation of the Heuristic

Inter-birth intervals are typically effected by parents' intended birth-spacing and the mother's fecundity. The former reflects parents' opportunity costs of time and resources, while the latter depends on factors such as maternal age, environmental conditions, post-partum amenorrhea, and contraceptive use. Sex-selection uniquely affects the length of the inter-birth interval *preceding* the birth at which sex-selection is employed.

Specifically, sex-selective abortion delays the next successful male birth, since fetal sex cannot be determined until around 10–12 weeks of gestation, and an abortion requires additional time for recovery before a new conception attempt can occur (Pörtner, 2022, 2016). The more attempts made to avoid female births, the longer the observed interval. On the other hand, if parents abandon sex selection after unsuccessful attempts, girls may also be born following extended intervals. The overall relationship between birth sex and the length of the preceding interval will then depend on the relative frequency of parents who abandon sex-selection versus those who persist. Likewise, if sex-selecting parents anticipate longer intervals and therefore initiate conception attempts earlier as a precaution, the resulting correlation may be ambiguous. Thus, any correlation between the sex of a newborn and the length of the preceding interval provides evidence of sex-selection.

Considering the temporal sequence of birth intervals and the legal revelation of a child's sex at birth[11]—that is, at the end of the interval; any observed correlation between the length of a birth interval and the sex of the child born at its conclusion implies a clear direction of effect: the sex outcome is influenced by parental behaviors and decisions made during that interval. Therefore, analyzing how the proportion of male births varies with the length of

[11]The Prenatal Sex Diagnostic Techniques (Regulation and Prevention of Misuse) Act of 1994 in India criminalizes the misuse of prenatal sex-diagnostic techniques to determine fetal sex, effectively banning prenatal sex screening.

the preceding birth interval offers valuable insights into the fertility choices parents make between successive live births.

Figure 3 plots the proportion of male births at the second birth order, conditional on the length of the preceding inter-birth interval. Separate subplots illustrate outcomes following different birth histories. If sex at consecutive births were determined independently, in the absence of sex selection, we would expect the plots to be identical across all birth histories and to remain flat at the natural probability of male births. Even if the relationship between the probability of a male birth and the length of the preceding inter-birth interval were nonlinear, independence in sex determination would still imply that the plots should, at a minimum, be consistent across different birth histories. Instead, Figure 3 suggests that mothers with firstborn girls are disproportionately more likely to have sons at the second birth, particularly following longer inter-birth intervals—a pattern consistent across the three NFHS rounds. Furthermore, the hump-shaped pattern suggests a reversal in the decision to sex-select when the inter-birth interval becomes "too long."

Similarly, Figures 4 and A7 plot the proportion of male births against the length of the preceding interval for children born at the third and fourth birth orders respectively. In both cases, for birth histories `FF` and `FFF`, the proportion of male births increases steadily with longer preceding intervals. Appendix Figures A1 -A6 repeat these comparisons at the second and third birth orders for different levels of mothers' education, places of residence, and geographical regions. The pattern exists across both urban and rural areas, as well as across all levels of maternal education. Comparisons by geographical region—also synonymous with cultural divisions in India—echo the well documented regional disparity in sex-ratios and gender relations in India, with strong sex-selecting behavior in the North-West and Central regions, contrasting with the South and North-East (Dyson & Moore, 1983).

I exploit this effect of the use of sex-selection on the preceding inter-birth interval to infer use of sex selection at a given birth order. At each birth order k, I compare correlations between sex of the child born to mother m and the interval preceding its birth (denoted by '$preI_{mk}$'), for various birth histories $h \in H_k$ referenced against the history with all male births. Here, H_k denotes the set of all sex-sequences until birth order k. For example, for births at the third birth order: $k = 3$ and $H_k = \{\texttt{MM, FF, MF, FM}\}$.

The following specification is used:

$$\begin{aligned} preI_{mk} = const + \sum_{h \in H_k \setminus h_r} \alpha_h \{\text{hist}_{mk}\text{:h}\} + \beta \{\text{sex}_{mk}\text{:M}\} + \\ \sum_{h \in H_k \setminus h_r} \gamma_h \{\text{hist}_{mk}\text{:h} \times \text{sex}_{mk}\text{:M}\} + \theta' X_{mk} + u_{mk} \end{aligned} \tag{3}$$

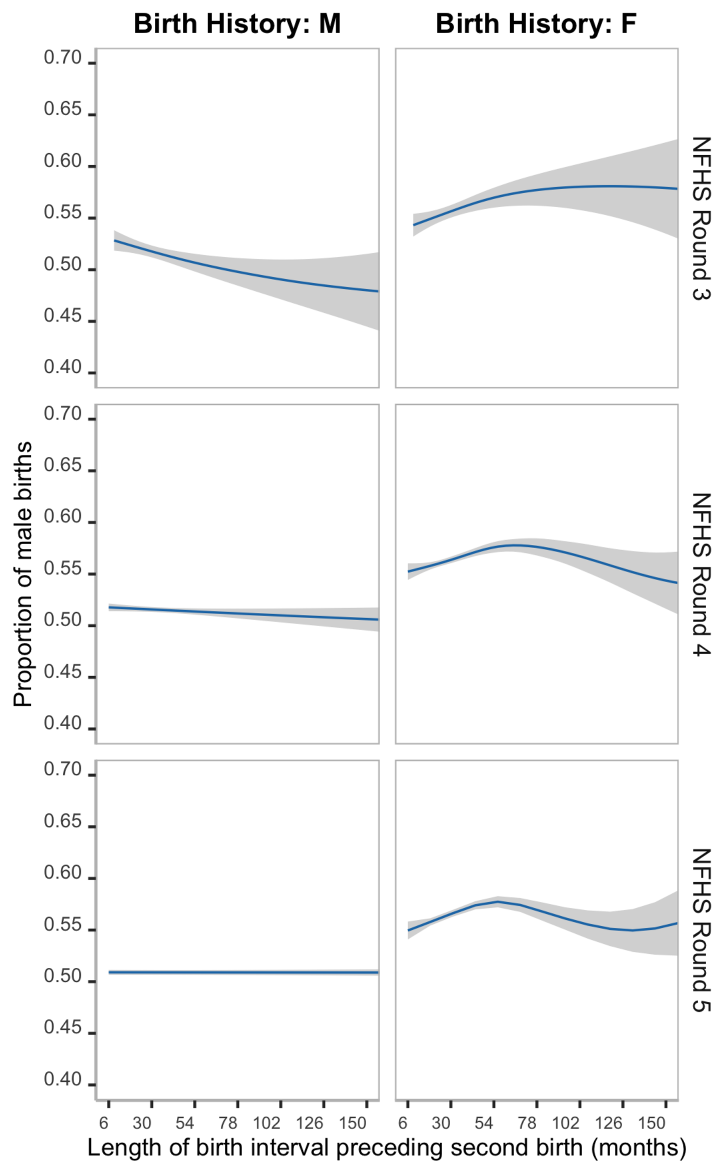
h_r : birth history with $k-1$ male children

hist_{mk}:h : 1 if birth history until order $k-1$ for mother m is h, 0 otherwise

sex_{mk}:M : 1 if k^{th} child of mother m is male, 0 otherwise

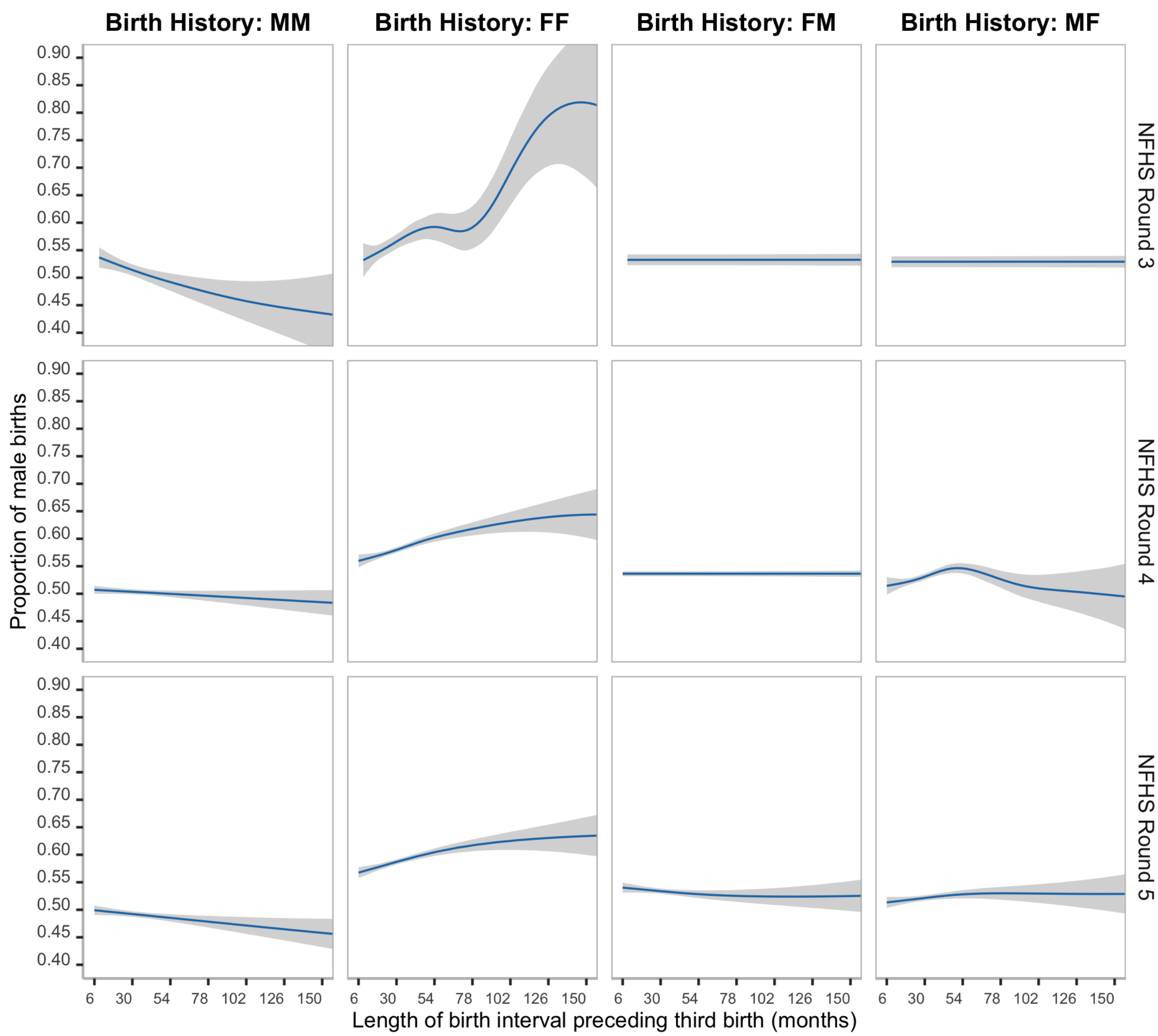
X_{mk} : vector of covariates

Figure 3 :
Proportion of male births (at the second birth order) at varying lengths of the preceding birth interval



Notes: This figure plots kernel-smoothed proportion of male births at the second birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories.

Figure 4 :
Proportion of male births (at the third birth order) at varying lengths of the preceding birth interval



Notes: This figure plots kernel-smoothed proportion of male births at the third birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories.

Equation (3) estimates the difference between intervals preceding male births to those preceding female births for a given history h, and then differences this difference relative to that for a reference history h_r with all male births. Here, the coefficients of interest are γ_h for all histories $h \in H_k \setminus h_r$, and significant coefficients indicate sex-selection. While the process of sex selection should produce a positive coefficient, reversals in the decision to sex-select following longer intervals may produce a negative coefficient. Nevertheless, a significant relationship between the child's sex and the length of the preceding birth interval—conditional on sex-selective histories—provides evidence of sex selection.

The identifying assumption here is that *even if* there are any unaccounted unobservables (such as geography, ethnicity, socio-economic or environmental factors) that could skew the natural probability of birthing a son *and* influence the length of the preceding interval; they should confound the relationship between sex and the length of the preceding interval, if at all, for *all* birth histories. Since the incentives to sex-select vary with birth history but the effect of confounding variables do not, any significant difference-in-difference in length of the preceding interval by the sex of the newborn across histories, must be attributed to parental manipulation.

Finally, I use this difference-in-difference specification to empirically validate the identified heuristic, as follows. For each birth order k, I divide mothers into two groups: those whose ideal parity preferences suggest they would sex-select at k based on the heuristic, and all others. The heuristic is considered validated if sex selection at birth order k—as captured by the difference-in-difference estimate—is predominantly driven by mothers in the first group.

IV. HEURISTIC OF SEX-SELECTION

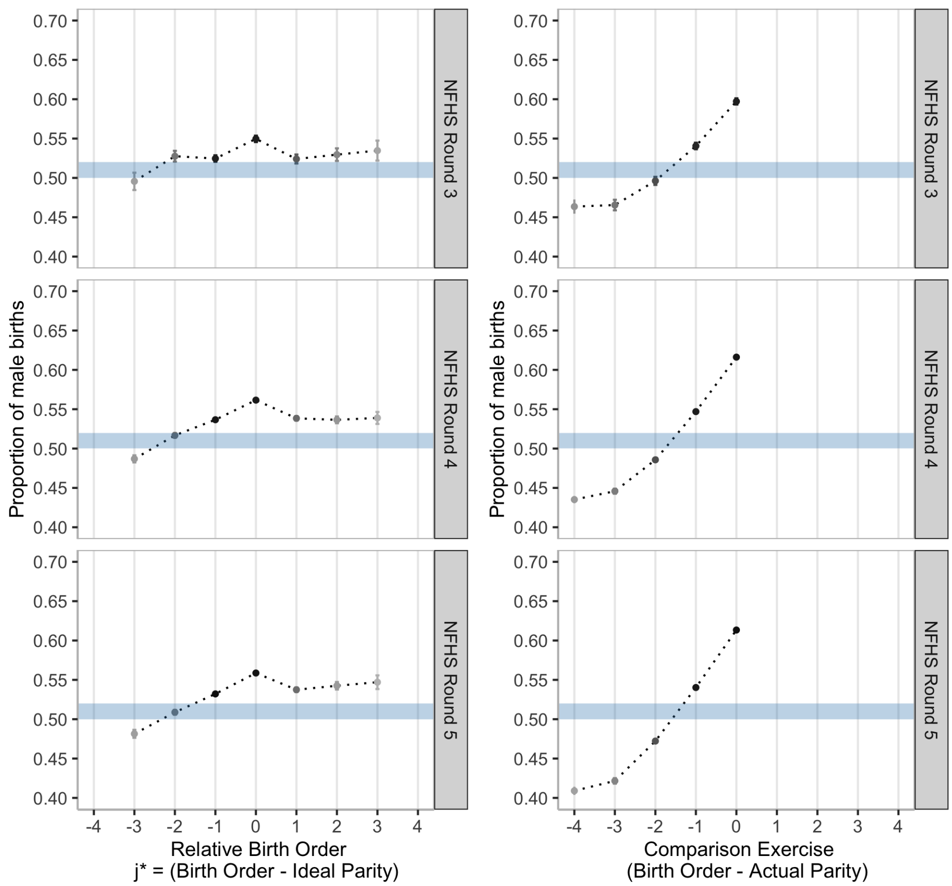
In this section, I identify a heuristic that broadly captures parents' decision-making around sex selection, and examine how it varies with parental characteristics, and provide empirical validation of the heuristic.

The heuristic is characterized by the relative birth orders that parents choose to sex-select at. The left panel of Figure 5 shows the proportion of sons born at each relative birth order, based on 2,231,621 births across the three NFHS rounds for mothers who report a strictly positive ideal parity.

The aggregate pattern shows a sharp increase in the proportion of male births at the relative birth order $j^* = 0$—that is, when mothers are at their ideal parity—followed by subsequent births at $j^* > 0$, where the proportion of male births remains slightly above but still close to the natural range. This aggregate pattern is, of course, shaped by the underlying composition of parents with varying fertility preferences and differing son-targeting strategies. Since parents who resort to sex selection would not subsequently switch to differential stopping, this pattern suggests the coexistence of two types of son-targeting behavior. The first group consists of exclusive sex-selectors, who sex-select at the relative birth order $j^* = 0$, that is, when they are at their ideal parity. The second group consists of differential stoppers, who exceed their ideal parity. Within this latter group, only a small fraction engage in sex selection, raising the proportion of male births at $j^* > 0$ slightly but not dramatically in the aggregate. This motivates the next step: examining how son-targeting behavior varies across different parental traits.

It is also useful to note that the increased tendency to stop childbearing after the birth of a son—whether through differential stopping or sex selection—is reflected in the high proportion of sons at last birth orders. This is illustrated in the right panel of Figure 5, which plots the proportion of sons at birth orders re-indexed based on the *distance from*

FIGURE 5 :
Proportion of male births by relative birth order



Notes: The left panel of this figure plots the proportion of male births at various relative birth orders. The right panel—as a comparison exercise—plots proportion of male births at birth orders re-indexed by the mother's actual parity. The shaded horizontal bar indicates the range 0.50-0.52 denoting the natural probability of male births Orzack et al. (2015). The gradient of the markers reflects the share of births contributing to the statistic at each relative birth order, with darker shades indicating higher representation.

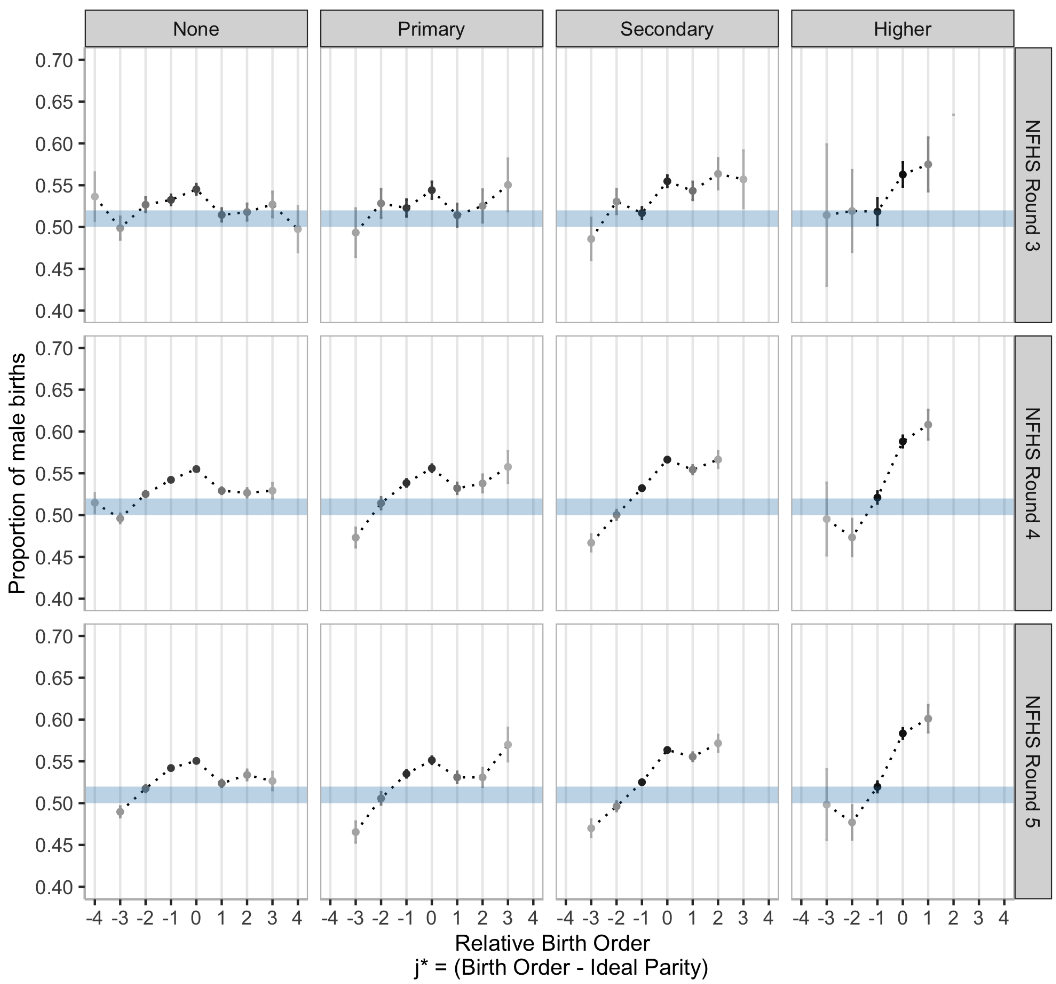
the actual parity instead, serving as a comparison exercise to the left panel. Notably, the dissimilarity between the two panels provides further evidence that respondents do not simply report their ideal parity to match their actual parity, lending credibility to my approach.

Variation in Heuristics by Parental Characteristics

Figure 6 examines how the heuristic of sex-selection varies with the level of mother's education. Two important patterns can be noted. First, mothers with higher levels of education ($\sim 7-8\%$ of the sample; Table B2) rarely exceed their ideal parity; and when they do, it is typically by only one child. This suggests that the cost of deviating from ideal parity likely increases with education. Second, mothers with higher levels of education exhibit a higher tendency to sex-select when at or exceeding their ideal parity. Sex selection is most commonly concentrated at $j^* = 0$ (e.g., $m_1 = 1, m_2 = 1$), and rarely at $j^* = 0$ and 1

$(m_1 = 1, m_2 = 2)$, or $j^* = 1$ $(m_1 = 0, m_2 = 1)$. In contrast, mothers with only primary or no education ($\sim 50-55\%$ of the sample; Table B2) often exceed their ideal parity when not sex-selecting $(m_1 < 0, m_2 = 0)$. Among those who do sex select, the practice typically occurs at or just before their stated ideal parity, most commonly at $j^* = -1, 0$ $(m_1 = 2, m_2 = 2)$ or at $j^* = 0$ $(m_1 = 1, m_2 = 1)$. The number of births that are sex selected (i.e. $m_2 = 1$ or 2) likely depends on how many sons are desired, which is once again negatively correlated with education. The reported ideal number of sons (likely under-reported) is mostly one in the sample, with nearly one-quarter of less-educated mothers reporting a desire for two sons; see Appendix Table B3.

FIGURE 6 :
Proportion of male births at relative birth orders, by mother's education



Notes: This figure plots the proportion of male births at various relative birth orders, by mother's education. The vertical columns represent the level of mother's education and the horizontal panels specify the NFHS round. The shaded horizontal bar indicates the range 0.50-0.52 denoting the natural probability of male births Orzack et al. (2015). The color gradient of the markers indicate the share of births in the sample contributing to the statistic at the relative birth order, with darker shades representing higher shares.

The increased tendency of well-educated mothers to engage in sex selection—also documented in Bhalotra & Cochrane (2010); Pörtner (2016, 2022)—stands in stark contrast to weakened son-preference typically expected with higher levels of education (Jayachandran, 2017). This apparent contradiction may be reconciled by the negative correlation between maternal education and ideal parity; see Appendix Table B4. Educated mothers are less willing to exceed their stated ideal parity unless sex selection guarantees a son (Table B5), effectively using sex selection as a controlled means to satisfy their fertility and sex composition preferences together.

A strikingly similar empirical pattern emerges when stratifying the sample by wealth quintile; see Appendix Figure A9. Moreover, the tendency to sex-select at the ideal parity—rather than exceed it—becomes increasingly prevalent across the wealth quintiles over time.

The fact that most parents resort to sex selection at or near their ideal number of children, regardless of the education level, is revealing: it underscores that parents who sex select are highly sensitive to achieving their desired family size without exceeding it. This matters because as desired parity declines over time, sex selection is likely to occur at earlier birth orders and be employed by more number of parents, since the probabilistic odds of achieving the desired number of sons within fewer births are low.

Appendix Figure A8 plots the proportion of sons born at various relative birth orders disaggregated jointly by geographical region and survey round. Consistent with the earlier patterns of sex selection varying by geographical region (Appendix Figures A3 and A6), Appendix Figure A8 shows clear and pronounced evidence of sex-selection around the ideal parity in states of the North-West, as well as in the Western and Central regions of India. In contrast, states in the South and East exhibit weak or no such patterns.

IV.A. Empirical Validation of the Heuristic

The findings so far indicate that most parents who engage in sex selection do so at $j^* = 0$. A subset of less-educated parents who sex-select twice typically do so at $j^* = -1$ and 0, while very few higher-educated parents exceed their ideal parity by one to sex-select at $j^* = 1$. Overall, sex selection at any birth order k is concentrated among parents whose ideal parity matches the birth order, i.e. for whom $n^* = k$. To empirically validate this, I first use the specification in Equation (3) to confirm sex selection at a given birth order k. Then, I validate the heuristic by showing that sex selection at birth order k is primarily driven by parents whose stated ideal parity is $n^* = k$. A good placebo sample to this exercise is mothers who one might otherwise want to sex-select but the heuristic doesn't expect them to. For births at any birth order k, this is effectively the sample of mothers with either $n^* < k - 1$ or with $n^* > k$.

Table 3 reports the estimates from the specification in Equation (3) for births at the second birth order. Table 3, column (1), provides evidence of sex selection at the second birth order: the coefficient on the interaction term indicates that boys are significantly more likely to be born after longer inter-birth intervals compared to girls, following a firstborn daughter (history `F`). Columns (2), (3) and (4) in Table 3 re-estimate the specification for three different sub-samples. Column (2) reports the estimates for mothers who the heuristic suggests are most likely to sex-select at the second birth order, i.e. mothers with $n^* = 2$.

Columns (3) and (4) repeat the exercise for placebo sample of mothers with $n^* > 2$ and $n^* < 1$ who we do not expect would resort to sex-selection at the second order. As expected, we find the coefficient of the interaction term 'hist: `F` × sex: `M`' to be significant in the sub-sample suggested by the heuristic, and statistically insignificant for the placebo sub-samples. The specification includes controls for the mother's age at previous birth, an indicator for rural residence, household wealth quintile, NFHS round, and state. Separately, the negative coefficient on 'hist : `F`' reaffirms the finding of parents beginning conception for the successive birth sooner after the birth of a female firstborn (Jayachandran & Kuziemko, 2011; Javed & Mughal, 2020).

Table 3 :
Empirical validation of the heuristic at the second birth order

	Dependent variable: Interval preceding **second** birth			
	All mothers with $N \geq 2$	Mothers with $n^* = 2$ (Heuristic)	Mothers with $n^* > 2$ (Placebo)	Mothers with $n^* < 1$ (Placebo)
	(1)	(2)	(3)	(4)
hist: F	-1.466^{***}	-1.297^{***}	-1.197^{***}	-1.898^{***}
	(0.071)	(0.099)	(0.104)	(0.415)
sex: M	-0.139^{**}	-0.309^{***}	0.065	-0.243
	(0.068)	(0.090)	(0.106)	(0.437)
hist: F × sex: M	0.550^{***}	0.610^{***}	-0.103	0.520
	(0.098)	(0.132)	(0.147)	(0.587)
Constant	29.809^{***}	28.833^{***}	34.815^{***}	33.851^{***}
	(0.245)	(0.316)	(0.462)	(1.739)
Observations	732,947	435,254	254,246	24,080
Note:			*p<0.1; **p<0.05; ***p<0.01	

This table reports the difference-in-difference estimates based on the specification in Equation (3) at the second birth order. Column (1) reports the estimates for all mothers who give birth to at least two children. Column (2) reports the estimates for mothers who according to our heuristic are expected to sex-select i.e. those whose ideal parity is equal to the given birth order ($n^* = 2$), and columns (3) and (4) do the same for placebo samples. The controls include mother's age at previous birth, rural, wealth quintile, NFHS round-specific fixed effects and state fixed effects. Robust standard errors included in parentheses.

Table 4 provides empirical support for the heuristic at the third birth order. Specifically, among mothers whose ideal parity is three, male births following two female births (history `FF`) are preceded by significantly longer inter-birth intervals—consistent with the use of sex selection. This pattern is not observed among mothers with other stated ideal parities, validating the heuristic.

TABLE 4 :
Empirical validation of the heuristic at the third birth order

	Dependent variable: Interval preceding **third** birth			
	All mothers with $N \geq 3$	Mothers with $n^* = 3$ (Heuristic)	Mothers with $n^* > 3$ (Placebo)	Mothers with $n^* < 2$ (Placebo)
	(1)	(2)	(3)	(4)
hist: FF	-1.347^{***}	-1.651^{***}	-1.758^{***}	-0.967
	(0.133)	(0.230)	(0.265)	(0.634)
hist: FM	-0.746^{***}	-0.463^{**}	-0.720^{***}	-1.190^{*}
	(0.132)	(0.219)	(0.268)	(0.656)
hist: MF	-1.238^{***}	-1.022^{***}	-1.149^{***}	-1.317^{**}
	(0.131)	(0.217)	(0.267)	(0.646)
sex: M	-0.463^{***}	-0.424^{**}	-0.129	-0.086
	(0.132)	(0.214)	(0.278)	(0.662)
hist: FF × sex: M	1.641^{***}	1.522^{***}	0.504	1.268
	(0.179)	(0.311)	(0.354)	(0.871)
hist: FM × sex: M	0.392^{**}	-0.026	-0.079	0.819
	(0.180)	(0.298)	(0.373)	(0.935)
hist: MF × sex: M	0.749^{***}	0.352	0.596	0.544
	(0.180)	(0.296)	(0.369)	(0.902)
Constant	28.206^{***}	27.713^{***}	30.806^{***}	32.228^{***}
	(0.339)	(0.635)	(0.870)	(1.782)
Observations	425,344	139,090	91,972	19,062
Note:			*p<0.1; **p<0.05; ***p<0.01	

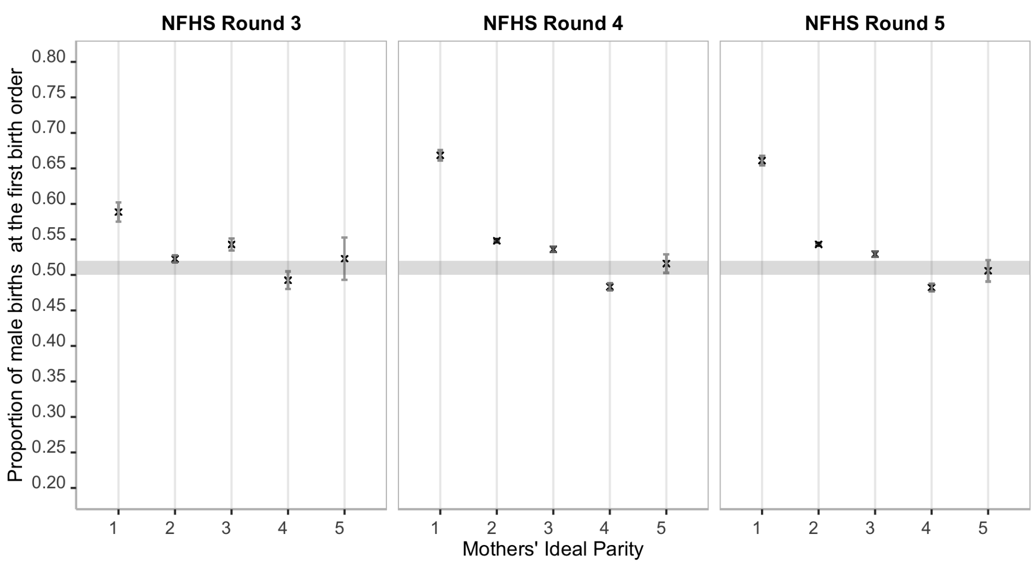
This table reports the difference-in-difference estimates based on the specification in Equation (3) at the third birth order. Column (1) reports the estimates for all mothers who give birth to at least three children. Column (2) reports the estimates for mothers who according to our heuristic are expected to sex-select i.e. those whose ideal parity is equal to the given birth order ($n^* = 3$), and columns (3) and (4) do the same for placebo samples. The controls include mother's age at previous birth, rural, wealth quintile, NFHS round-specific fixed effects and state fixed effects. Robust standard errors included in parentheses.

Caution is warranted when interpreting results at higher birth orders, as the sample size becomes increasingly limited. Additionally, analyses at $k \geq 4$ inherently focus on a selective group of parents—those willing to have at least k children. As k increases, this group is increasingly composed of parents with either a high ideal parity or a greater willingness to exceed their stated fertility preferences. This selection limits the generalizability of findings at higher parities. Appendix Table B6 reports the estimates for the fourth birth order. We find evidence of sex-selection following histories `FFF`, `FMF`, and `MFF`. At the fourth order, sex-selection is not driven by mothers reporting an ideal parity of four, but rather by parents whose preferred family size is smaller but continued childbearing to later sex-select, or by parents now seeking to balance the sexes. Moreover, the sample size underlying the results for specific histories at higher birth orders also declines sharply.

IV.B. Sex-Selection at the First Birth-Order

Previous studies have typically assumed that sex selection does not occur at the first birth order, citing aggregate sex ratios at the first birth order to fall within the "normal" range. However, when disaggregated by mothers' stated ideal parity, a different pattern emerges: mothers who report an ideal parity of one are disproportionately more likely to have a male firstborn (Figure 7). Moreover, for this group, the proportion of male births increases with the length of the interval between marriage and first birth (Figure 8). No such trend is observed among mothers with higher stated ideal parities. This calls for a closer examination of what happens at the first birth order. However, since there are no prior birth histories before the first birth, the specification in Equation (3) cannot be applied. Additionally, the interval preceding the first birth—the time from marital union to first birth—is also conceptually distinct from inter-birth intervals.

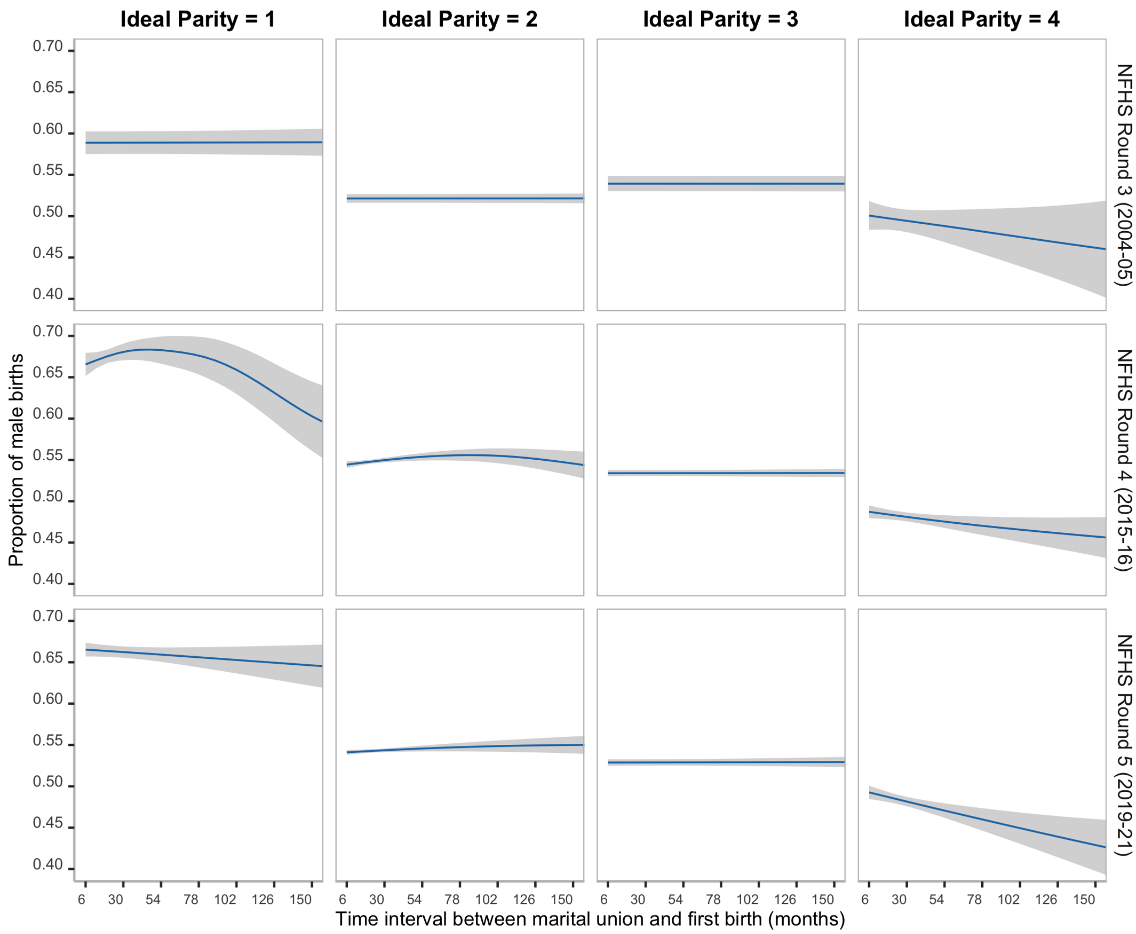
FIGURE 7 :
Proportion of male births at the first birth order



Notes: This figure plots the proportion of male births at the first birth order, disaggregated by mothers' ideal parity and the NFHS round. The shaded horizontal bar indicates the range 0.50-0.52 denoting the natural probability of male births Orzack et al. (2015).

FIGURE 8 :

Proportion of male births (at the first birth order) at varying lengths of the preceding birth interval



Notes: This figure plots kernel-smoothed proportion of male births at the first birth order, conditional on the length of preceding interval between marital union and first birth. Each panel plots the conditional proportions for mothers with different values of ideal parity.

To address these challenges, I draw on the heuristic identified earlier. According to the heuristic, parents most likely to sex-select at the first birth order are those with an ideal parity of one ($n^* = 1$). I implement a difference-in-differences design that compares the length of the interval from marital union to the first birth, by the sex of the firstborn, across sub-samples predicted by the heuristic to sex-select versus those not predicted to do so. Specifically, I contrast mothers with an ideal parity of one ($n^* = 1$)—the group likely to sex-select at the first birth order—with a placebo sub-sample of mothers whose ideal parity exceeds one.

The estimating equation is:

$$\begin{aligned} log(I_{m1}) = const + \alpha \mathbb{1}\{\text{sub-sample qualifier}\} + \beta\{\text{sex:M}\} + \\ \gamma \mathbb{1}\{\text{sub-sample qualifier}\} \times \{\text{sex:M}\} + \theta' X_{m1} + u_{m1} \end{aligned} \tag{4}$$

$$\begin{aligned} m &: \text{index for mother} \\ I_m &: \text{interval between marriage/union and the first birth for } m \\ \text{sex:M} &= 1 \text{ if first child is male, 0 otherwise} \\ \text{sub-sample qualifier} &: \mathbb{1}\{n^* = 1\} \text{ for the heuristic, and } \mathbb{1}\{n^* > 1\} \text{ as placebo} \\ X_m &: \text{vector of covariates} \end{aligned}$$

If the coefficient γ is statistically significant for mothers with ideal parity of one, but insignificant for the placebo group, this provides evidence of sex selection at the first birth order. Table 5 reports the estimates for the specification. Column (1) reports the correlation for all mothers in the sample. Columns (2) and (3) report the estimates for the heuristic and placebo sub-samples. Mothers with an ideal parity of one exhibit longer intervals between marital union and first birth relative to other mothers who report higher ideal parities. Within this group, however, mothers who give birth to boys at the first birth order tend to conceive significantly sooner, consistent with some parents initiating conception attempts earlier in anticipation of delays from sex selection. By contrast, among mothers with preferred parity greater than one, there is no systematic relationship between the sex of the firstborn and the timing of the first birth.

V. CONCLUSION AND DISCUSSION

The prevalence of son preference and the associated practices of differential stopping and sex selection are well documented in the Indian context. This paper contributes by characterizing parental choice between the two son-targeting methods—particularly, identifying when parents opt to engage in sex selection. As a general heuristic summarizing the behavior of most Indian parents, I find that parents are more likely to sex-select when at their ideal parity in order to avoid exceeding it.

This finding is particularly pertinent in light of the declining trend in desired family size. As ideal parity falls, parents are more likely to sex-select at lower birth orders. Since achieving a preferred sex composition becomes increasingly difficult with fewer births, this shift implies not only a move from higher to lower parity sex-selection, but also a potential increase in the overall incidence of sex-selection. Consistent with this, I find evidence of sex-selection even at the first birth.

The heuristic of sex-selecting at ideal parity is consistently observed across both rural and urban areas, and across varying levels of maternal education and household wealth. In addition to this general pattern, more educated and wealthier parents also exhibit a tendency to exceed their ideal parity by at most one child, who is likely sex-selected. The broad consistency of this behavior across socioeconomic groups cautions against overly simplistic assumptions that sex-selection will naturally decline with economic development or improvements in maternal education.

TABLE 5 :
Sex-selection at the first birth order, using the heuristic

	Dependent variable: Interval preceding **first** birth (since marriage/union)		
	All	Heuristic ($n^* = 1$)	Placebo ($n^* > 1$)
	(1)	(2)	(3)
sex: M	-0.103^{*}	-0.201^{***}	-0.303
	(0.053)	(0.054)	(0.230)
Heuristic: $\mathbb{1}\{n^* = 1\}$		6.698^{***}	
		(0.258)	
Placebo: $\mathbb{1}\{n^* > 1\}$			-4.787^{***}
			(0.188)
sex M $\times$ Heuristic: $\mathbb{1}\{n^* = 1\}$		-1.073^{***}	
		(0.309)	
sex M $\times$ Placebo: $\mathbb{1}\{n^* > 1\}$			0.014
			(0.236)
Constant	56.126^{***}	56.629^{***}	61.203^{***}
	(0.289)	(0.291)	(0.352)
Observations	742,362	737,404	737,404
Note:			*p<0.1; **p<0.05; ***p<0.01

This table reports the difference-in-difference estimates based on the specification in Equation (4) at the first birth order. Column (1) reports the estimates for all mothers in our sample. Columns (2) and (3) report the estimates using the sub-sample qualifiers for the heuristic and placebo, respectively. The controls include mother's age at previous birth, rural, wealth quintile, NFHS round-specific fixed effects and state fixed effects. Robust standard errors included in parentheses.

Following the limited effectiveness of the Pre-Conception and Pre-Natal Diagnostic Techniques (PCPNDT) Act in curbing sex-selection, the Indian government has experimented with financial incentives to encourage greater valuation of daughters. These efforts have had mixed outcomes: while the Devi Rupak program in Haryana reportedly backfired (Anukriti et al., 2022), the Dhanlakshmi scheme showed some success in influencing parental attitudes (Biswas et al., 2023). However, such financial interventions are both fiscally costly and likely to influence behavior primarily only among lower-income households (Jayachandran, 2024).

Having identified a general heuristic for when parents are likely to engage in sex selection, we can consider potential policy interventions that more precisely target such behavior. A key challenge, however, lies in the lack of reliable data on desired parity preferences. One possible approach is to leverage local Accredited Social Health Activists (ASHAs) or Auxiliary Nurse Midwives (ANMs) to engage with women of reproductive age and gather information on their fertility preferences, while recognizing the possibility that these actors could themselves be assisting women or families in their communities in engaging in sex selection. Another complementary measure could involve conducting random audits at delivery clinics or hospitals to detect skewed sex ratios at birth, particularly at lower birth orders.

Addressing differential stopping through policy is more challenging, as directly curbing fertility choices raises ethical concerns, and reducing son preference remains a complex and long-term endeavor. A more pragmatic policy approach may lie in mitigating the adverse consequences of differential stopping by addressing gender disparities in nutritional and educational investments—through universal child welfare programs as well as targeted initiatives focused on young women's' skill development and employment opportunities. value differential between sons and daughters. In the longer term, sustained efforts should aim to eliminate the perceived value differential between sons and daughters.

References

Anderson, S., & Ray, D. (2010). Missing Women: Age and Disease. *Review of Economic Studies*, *77*, 1262-1300.

Anukriti, S., Bhalotra, S., & Tam, E. H. (2022). On the Quantity and Quality of Girls: Fertility, Parental Investments, and Mortality. *The Economic Journal*, *132*.

Baland, J.-M., Cassan, G., & Woitrin, F. (2020). The Stopping Rule and Gender selective mortality: World Evidence. *CEPR Discussion Papers*.

Baland, J.-M., Cassan, G., & Woitrin, F. (2023). Sex-Selective Abortions and Instrumental Births as the two faces of the Stopping Rule. New measures and World Evidence. *CEPR Discussion Papers*.

Basu, D., & Jong, R. D. (2010). Son targeting fertility behavior: Some consequences and determinants. *Demography*, *47*, 521-536.

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

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

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 of the United States of America*, *116*, 9303–9311.

Das, N. (1987). Sex Preference and Fertility Behavior: A Study of Recent India Data. *Demography*, 517-530.

Dyson, T., & Moore, M. (1983). On Kinship Structure, Female Autonomy, and Demographic Behavior in India. *Population and Development Review*, 35-60. doi: 10.2307/1972894

Javed, R., & Mughal, M. (2020). Preference for Boys and Length of Birth Intervals in Pakistan. *Research in Economics*, 140-152.

Jayachandran, S. (2015). The Roots of Gender Inequality in Developing Countries. *Annual Review of Economics*, 63-88.

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

Jayachandran, S. (2024). Ten facts about son preference in india. *India Policy Forum, 2023*.

Jayachandran, S., & Kuziemko, I. (2011). Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India. *The Quarterly Journal of Economics*, *126*, 1485-1538.

Jayachandran, S., & Pande, R. (2017). Why Are Indian Children So Short? The Role of Birth Order and Son Preference. *American Economic Review*, 2600-2629.

Jha, P., Kumar, R., Vasa, P., Dhingra, N., Thiruchelvam, D., & Moineddin, R. (2006). Low Male-to-Female Sex Ratio of Children Born in India: National Survey of 1· 1 Million Households. *The Lancet*, *367*, 211-218.

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

Pande, R. P., & Astone, N. M. (2007). Explaining Son Preference in Rural India: The Independent Role of Structural versus Individual Factors. *Population Research and Policy Review*, *26*(1), 1-29.

Pörtner, C. C. (2016). Sex-Selective Abortions, Fertility and Birth Spacing. *World Bank Policy Research Working Paper, no. WPS 7189.*.

Pörtner, C. C. (2022). Birth Spacing and Fertility in the Presence of Son Preference and Sex-Selective Abortions: India's Experience Over Four Decades. *Demography*, *59*(1), 61-88.

Rasul, I. (2008). Household Bargaining Over Fertility: Theory and Evidence From Malaysia. *Journal of Development Economics*, 215-41.

Sen, A. (1990). More Than 100 Million Women Are Missing. *The New York Review of Books*.

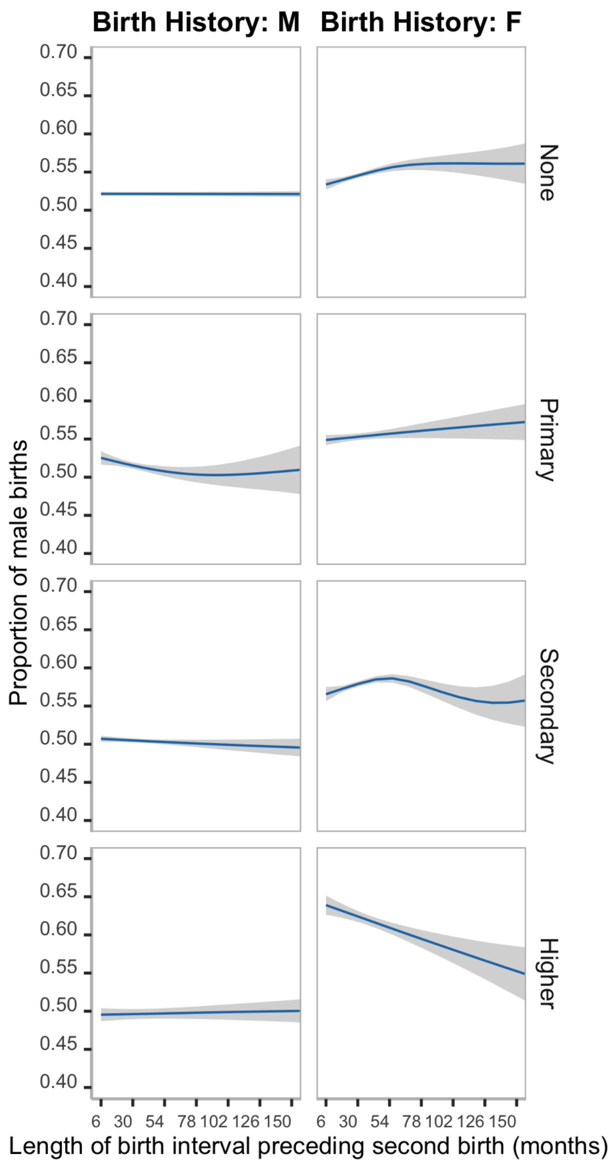
Sen, A. (1992). Missing Women. *British Medical Journal*, *304*, 587–588.

Tables created in R using stargazer v.5.2.2 by Marek Hlavac, Harvard University.

APPENDIX

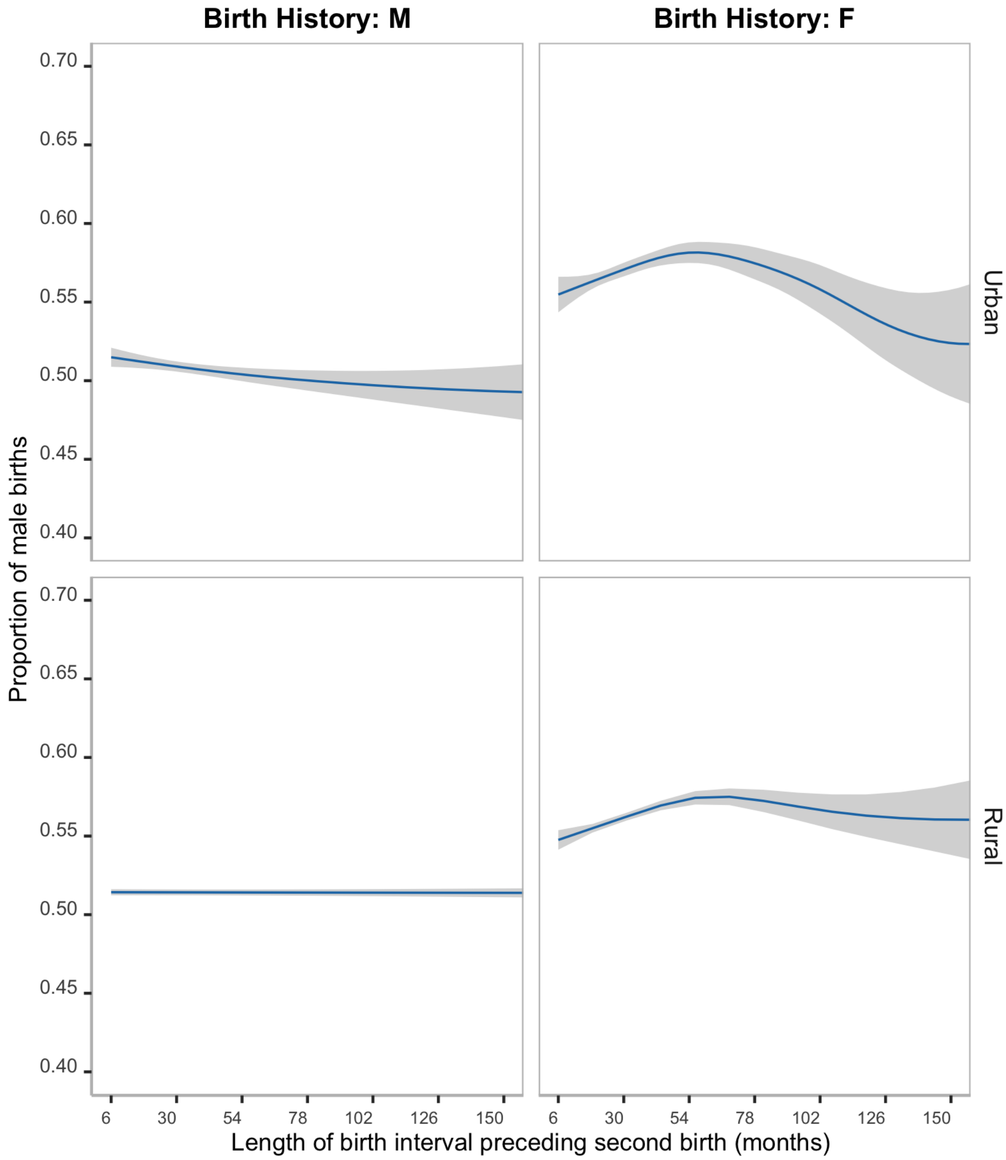
A. FIGURES

Figure A1 :
Proportion of male births (at the second birth order) at varying lengths of the preceding birth interval, by mother's education



Notes: This figure plots kernel-smoothed proportion of male births at the second birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories and across various levels of mother's education.

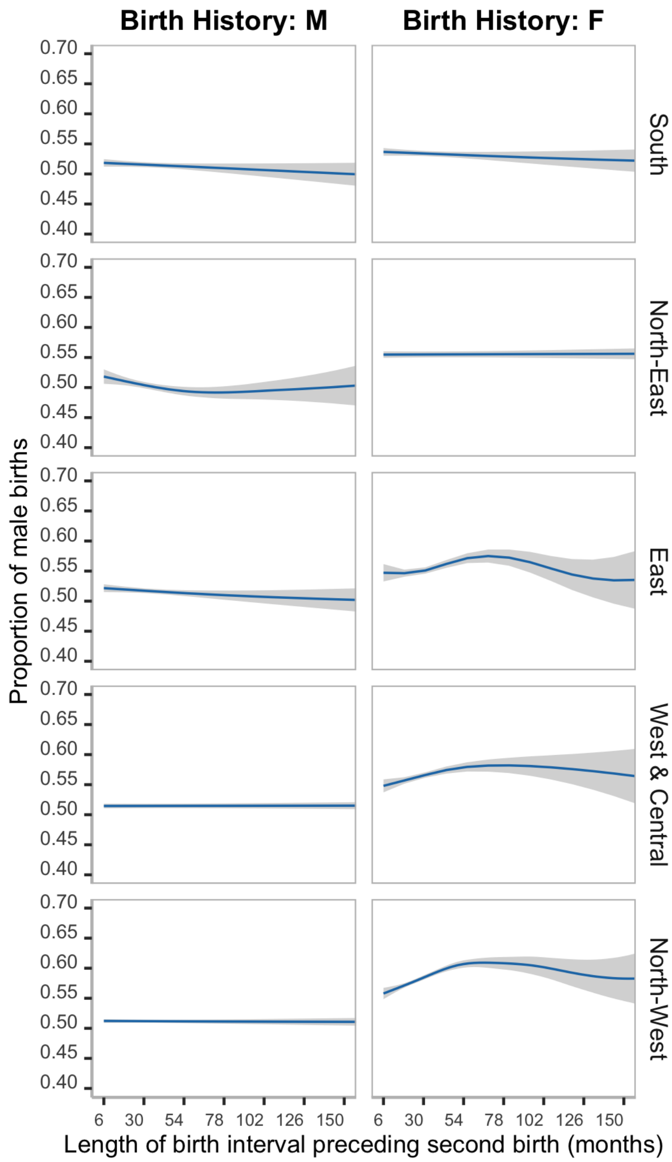
Figure A2 :
Proportion of male births (at the second birth order) at varying lengths of the preceding birth interval, by place of residence



Notes: This figure plots kernel-smoothed proportion of male births at the second birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories and across urban/rural settings.

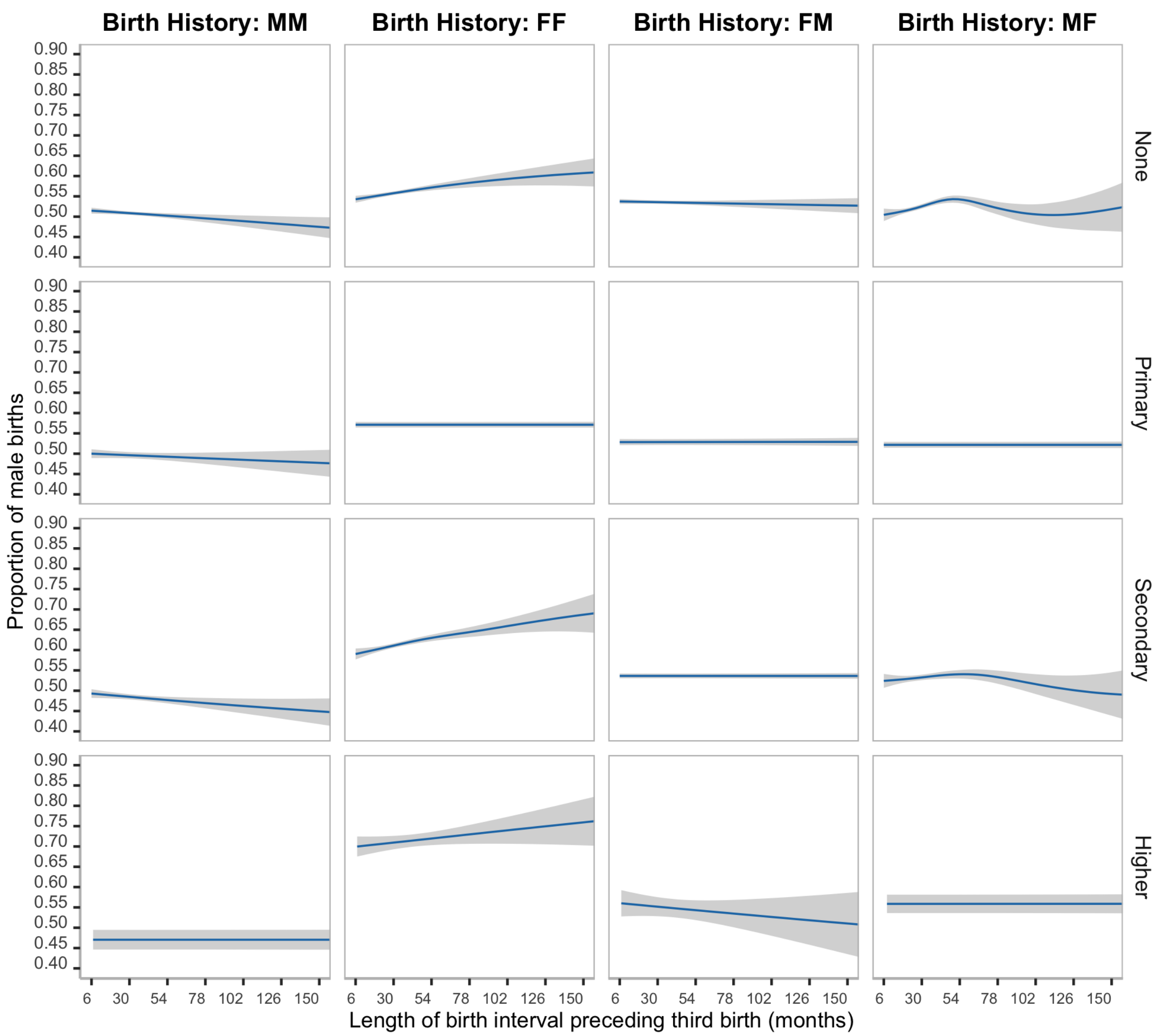
Figure A3 :

Proportion of male births (at the second birth order) at varying lengths of the preceding birth interval, by geographical region



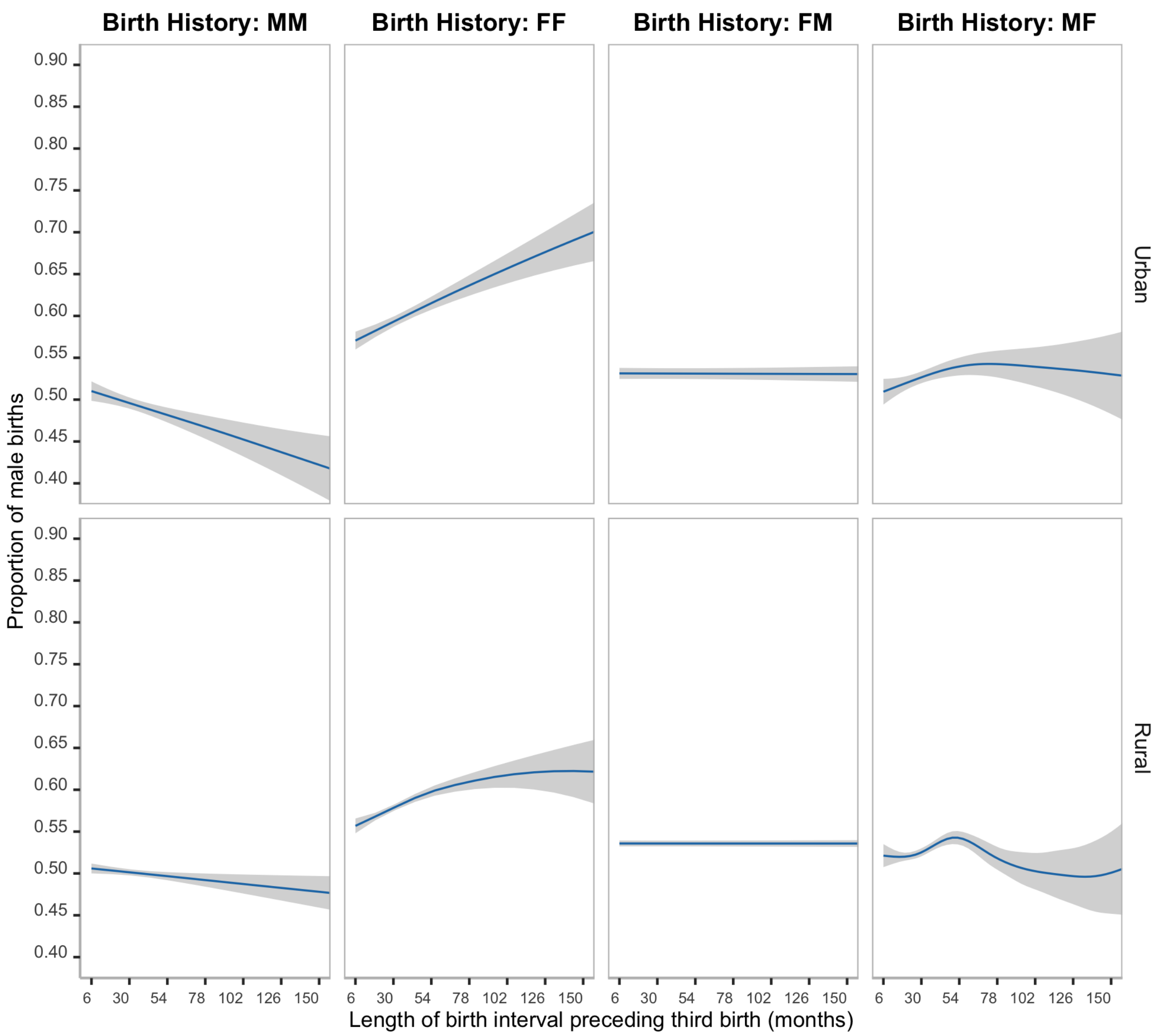
Notes: This figure plots kernel-smoothed proportion of male births at the second birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories and across various geographical regions in India.

Figure A4 :
Proportion of male births (at the third birth order) at varying lengths of the preceding birth interval, by mother's education



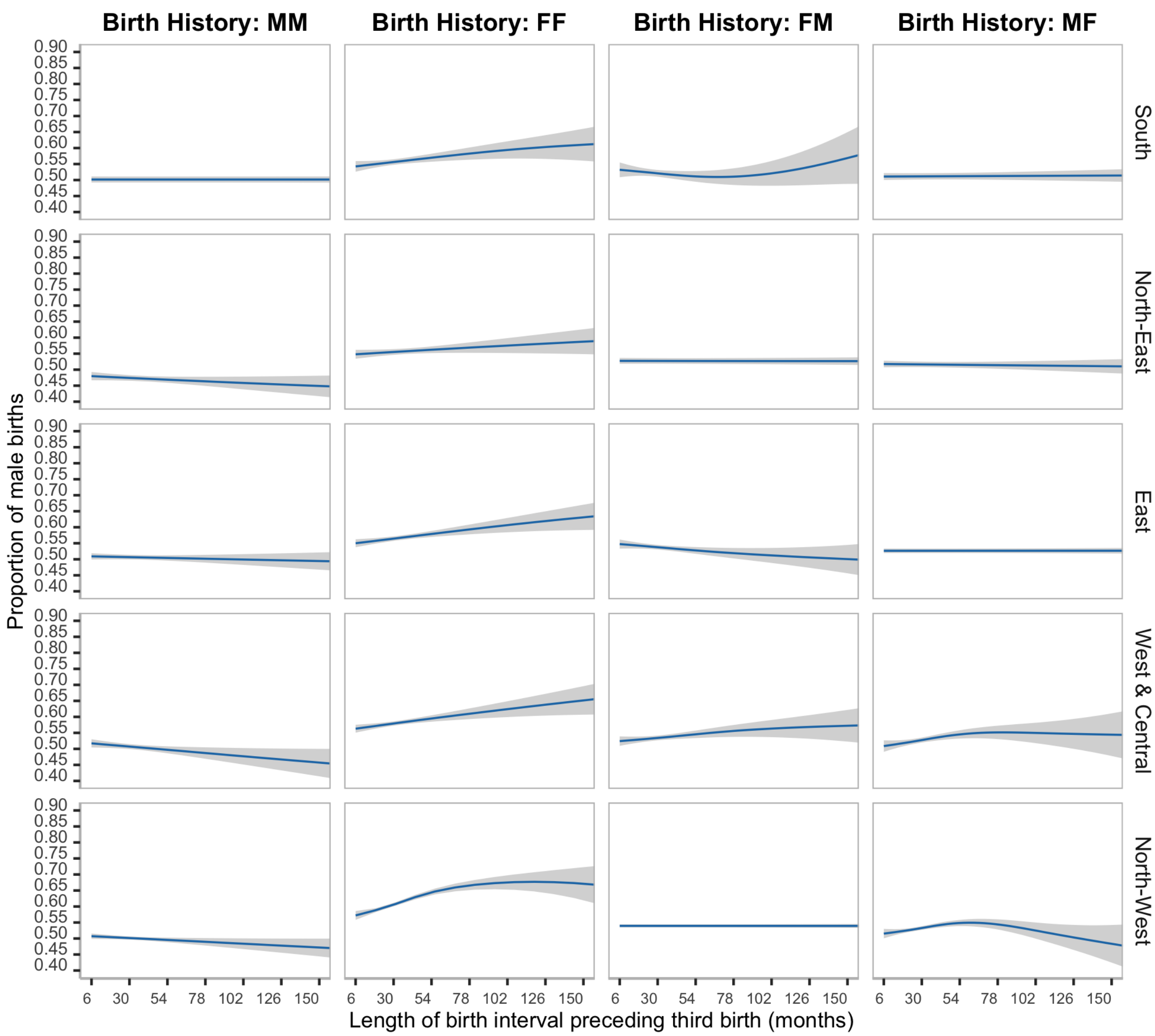
Notes: This figure plots kernel-smoothed proportion of male births at the third birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories and across various levels of mother's education.

Figure A5 :
Proportion of male births (at the third birth order) at varying lengths of the preceding birth interval, by place of residence



Notes: This figure plots kernel-smoothed proportion of male births at the third birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories and across urban/rural settings.

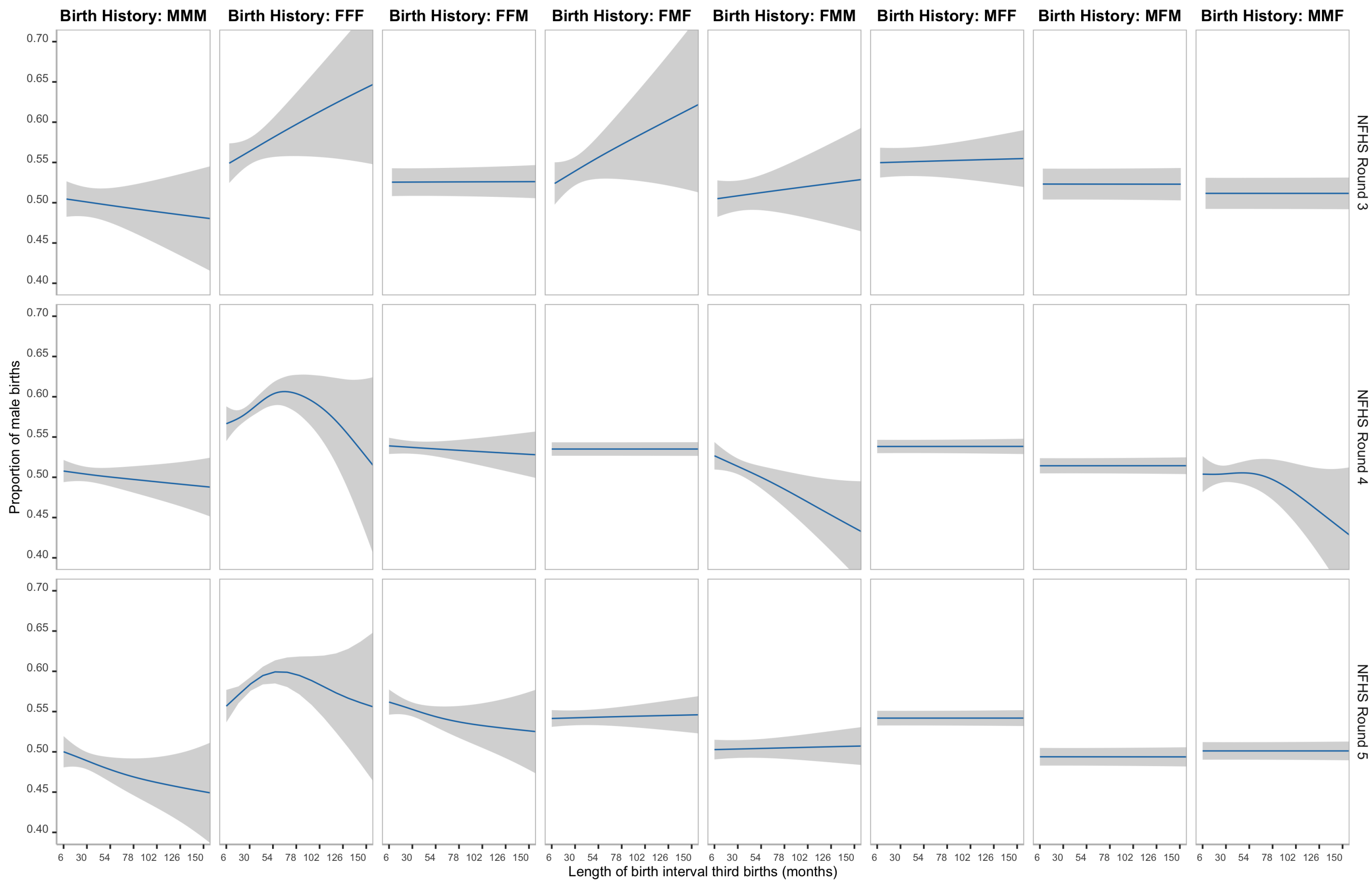
Figure A6 :
Proportion of male births (at the third birth order) at varying lengths of the preceding birth interval, by geographical region



Notes: This figure plots kernel-smoothed proportion of male births at the third birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories and across various geographical regions in India.

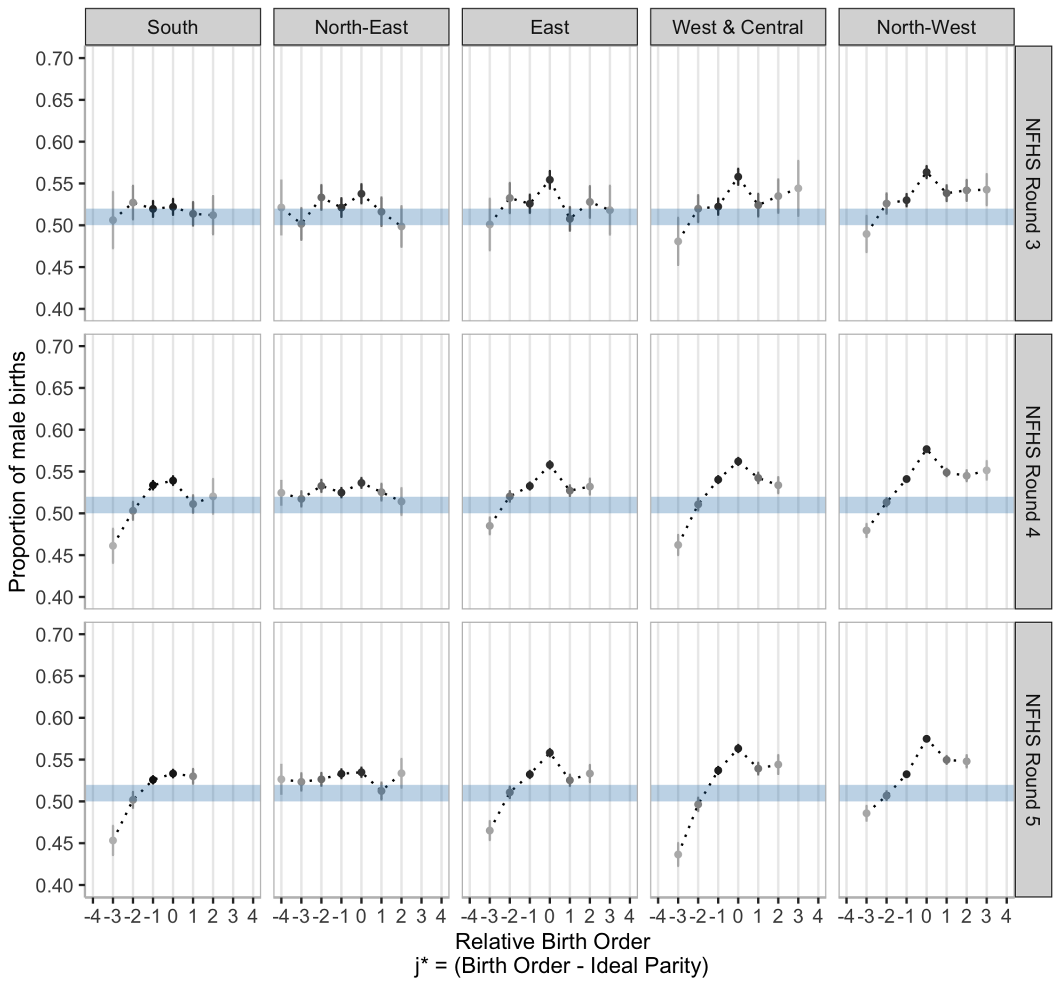
Figure A7 :

Proportion of male births (at the fourth birth order) at varying lengths of the preceding inter-birth interval



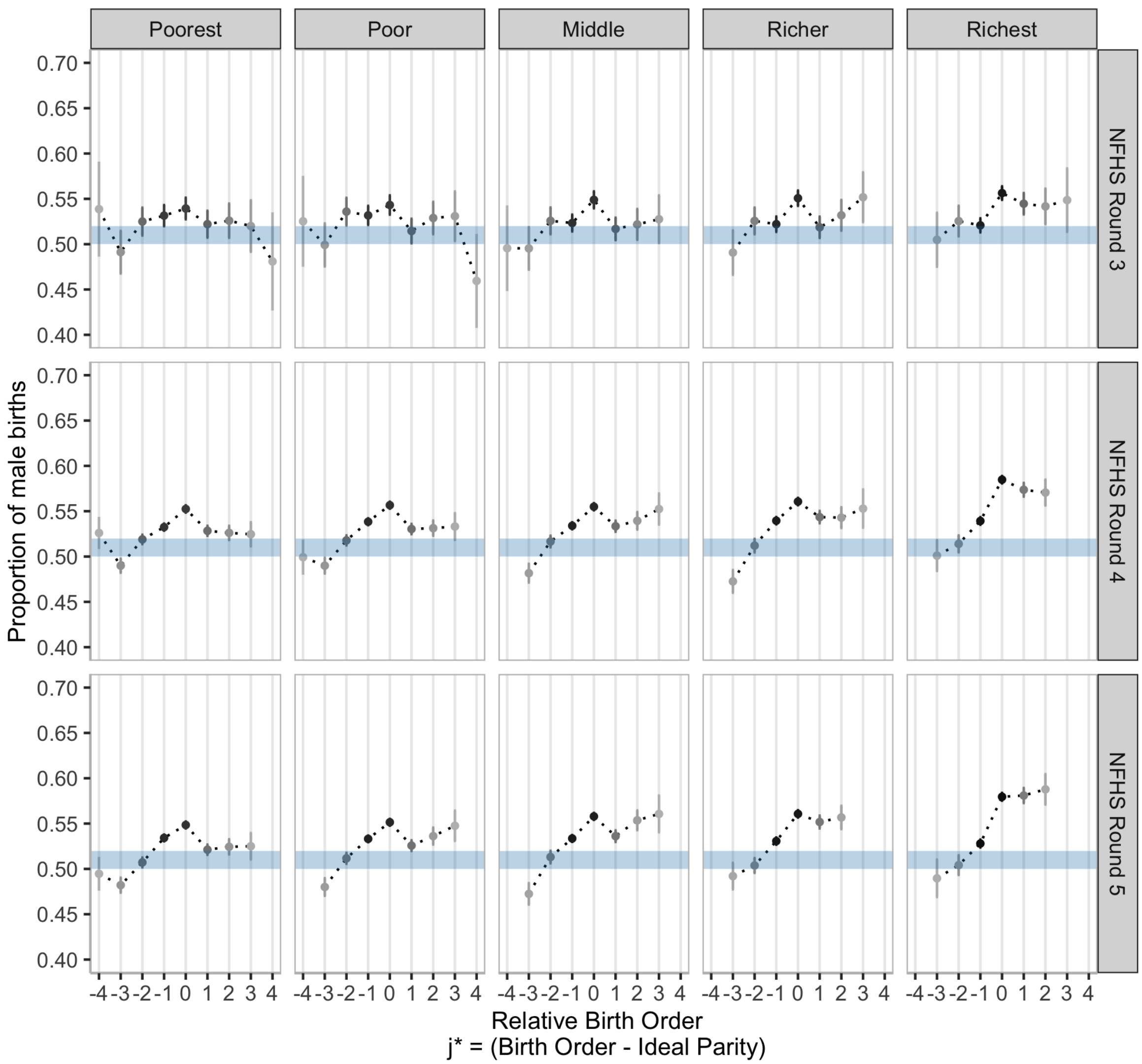
Notes: This figure plots kernel-smoothed proportion of male births at the fourth birth order, conditional on the length of the preceding interval. Each panel plots the conditional proportions following different birth histories.

Figure A8 :
Proportion of male births at relative birth orders, by geographical region



Notes: This figure plots the proportion of male births at various relative birth orders, by geographical regions. The vertical columns represent the geographical region and the horizontal panels specify the NFHS round. The shaded horizontal bar indicates the range 0.50-0.52 denoting the natural probability of male births Orzack et al. (2015). The color gradient of the markers indicate the share of births in the sample contributing to the statistic at the relative birth order, with darker shades representing higher shares.

Figure A9 :
Proportion of male births at relative birth orders, by parents' wealth quintile



Notes: This figure plots the proportion of male births at various relative birth orders, by the parents' wealth quintile. The vertical columns represent the five wealth quintiles and the horizontal panels specify the NFHS round. The shaded horizontal bar indicates the range 0.50-0.52 denoting the natural probability of male births Orzack et al. (2015). The color gradient of the markers indicate the share of births in the sample contributing to the statistic at the relative birth order, with darker shades representing higher shares.

B. TABLES

TABLE B1 :
Correlation between ideal and actual parity

	Dependent variable: Ideal parity		
	NFHS-3 (2005-06)	NFHS-4 (2015-16)	NFHS-5 (2019-21)
Total parity	0.290***	0.356***	0.362***
	(0.002)	(0.001)	(0.001)
$\mathbb{1}$(multiple births)	-0.019	0.064**	0.020
	(0.056)	(0.027)	(0.025)
Total parity $\times \mathbb{1}$(multiple births)	−0.040***	−0.064***	−0.059***
	(0.010)	(0.005)	(0.006)
Constant	1.455***	1.329***	1.314***
	(0.007)	(0.003)	(0.003)
Observations	67,445	371,869	379,831
R^2	0.261	0.256	0.256
Note:			*p<0.1; **p<0.05; ***p<0.01

Notes: This table examines the correlation between mothers' report ideal parities and their actual realized parity. In particular, the correlation is low and reduces in the case of multiple births i.e. when a pregnancy produces more than one baby (e.g. twins, triplets).

Table B2 :
Maternal Education (Distribution)

	Level of Maternal Education			
NFHS Round	None	Primary	Secondary	Higher
3	0.4172	0.1612	0.3449	0.0765
4	0.4017	0.1568	0.3767	0.0648
5	0.3414	0.1575	0.4224	0.0786

Table B3 :
Reported Ideal Number of Sons (Distribution)

	Reported Ideal Number of Sons (b^*)				
Level of Maternal Education	0	1	2	3	4
None	0.1060	0.4447	0.3999	0.0399	0.0071
Primary	0.1233	0.5838	0.2660	0.0220	0.0035
Secondary	0.1615	0.6620	0.1637	0.0110	0.0014
Higher	0.2502	0.6751	0.0714	0.0029	0.0003

Table B4 :
Ideal Parity: Correlation with Maternal Education

	Dependent variable: Ideal Parity		
	OLS		Poisson
	(1)	(2)	(3)
edu: Primary	-0.238^{***}	-0.209^{***}	-0.083^{***}
	(0.003)	(0.003)	(0.002)
edu: Secondary	-0.466^{***}	-0.404^{***}	-0.169^{***}
	(0.002)	(0.003)	(0.002)
edu: Higher	-0.699^{***}	-0.601^{***}	-0.268^{***}
	(0.004)	(0.005)	(0.004)
Constant	2.386^{***}	2.519^{***}	0.908^{***}
	(0.007)	(0.008)	(0.006)
Observations	802,958	802,958	802,958
Controls		X	X
Adjusted R^2	0.190	0.195	
Note:			*p<0.1; **p<0.05; ***p<0.01

Notes: This table examines the correlation between mothers' report ideal parities and their level of education. The controls include wealth quintile, urban/rural indicator, NFHS round-specific fixed effects and state fixed effects. Standard errors included in parentheses.

Table B5 :
Exceeding Ideal Parity: Correlation with Maternal Education

	Dependent variable: Total parity exceeds ideal parity		
	Linear Probability Model		Logistic
	(1)	(2)	(3)
edu: Primary	-0.067^{***}	-0.062^{***}	-0.243^{***}
	(0.002)	(0.002)	(0.007)
edu: Secondary	-0.193^{***}	-0.179^{***}	-0.771^{***}
	(0.001)	(0.001)	(0.006)
edu: Higher	-0.342^{***}	-0.313^{***}	-1.595^{***}
	(0.002)	(0.002)	(0.013)
Constant	0.500^{***}	0.530^{***}	0.121^{***}
	(0.004)	(0.004)	(0.020)
Observations	802,958	802,958	802,958
Controls		X	X
Adjusted R^2	0.085	0.086	
Note:			*p<0.1; **p<0.05; ***p<0.01

Notes: This table examines the relationship between mothers' education levels and their likelihood of exceeding their stated ideal parity. The controls include wealth quintile, urban/rural indicator, NFHS round-specific fixed effects and state fixed effects. Standard errors included in parentheses.

Table B6 :
Empirical validation of the heuristic at the fourth birth order

	Dependent variable: Interval preceding **fourth** birth			
	All mothers with $N \geq 4$	Mothers with $n^* = 4$ (Heuristic)	Mothers with $n^* > 4$ (Placebo)	Mothers with $n^* < 3$ (Placebo)
	(1)	(2)	(3)	(4)
hist: FFF	-1.413^{***}	-1.259^{***}	-1.644^{**}	-1.866^{***}
	(0.241)	(0.442)	(0.773)	(0.452)
hist: FFM	-0.908^{***}	0.039	-1.035	-1.945^{***}
	(0.245)	(0.440)	(0.801)	(0.470)
hist: FMF	-1.745^{***}	-0.970^{**}	-2.498^{***}	-2.599^{***}
	(0.250)	(0.450)	(0.782)	(0.479)
hist: FMM	-0.367	0.549	-1.281	-1.036^{*}
	(0.272)	(0.471)	(0.828)	(0.556)
hist: MFF	-1.268^{***}	-0.372	-0.963	-2.446^{***}
	(0.250)	(0.458)	(0.797)	(0.477)
hist: MFM	-0.642^{**}	0.302	-1.369^{*}	-1.424^{***}
	(0.271)	(0.470)	(0.808)	(0.547)
hist: MMF	-0.486^{*}	0.266	-1.425^{*}	-0.758
	(0.274)	(0.464)	(0.817)	(0.564)
sex: M	-0.708^{***}	0.376	-0.528	-1.767^{***}
	(0.263)	(0.475)	(0.812)	(0.518)
hist: FFF × sex: M	1.387^{***}	0.042	0.179	2.748^{***}
	(0.329)	(0.619)	(1.077)	(0.607)
hist: FFM × sex: M	0.293	-1.029^{*}	0.739	1.252^{**}
	(0.335)	(0.613)	(1.116)	(0.631)
hist: FMF × sex: M	1.027^{***}	0.094	3.034^{***}	1.627^{**}
	(0.344)	(0.630)	(1.119)	(0.646)
hist: FMM × sex: M	0.624	-0.323	1.451	0.870
	(0.389)	(0.740)	(1.156)	(0.746)
hist: MFF × sex: M	0.895^{**}	-0.261	-0.894	2.085^{***}
	(0.349)	(0.632)	(1.098)	(0.646)
hist: MFM × sex: M	0.686^{*}	-0.700	0.518	1.586^{**}
	(0.376)	(0.676)	(1.132)	(0.743)
hist: MMF × sex: M	0.565	-0.314	2.058^{*}	0.916
	(0.378)	(0.675)	(1.147)	(0.758)
Constant	28.672^{***}	28.107^{***}	30.581^{***}	29.853^{***}
	(0.552)	(1.069)	(2.695)	(0.907)
Observations	214,529	62,176	16,017	72,182
Note:				*p<0.1; **p<0.05; ***p<0.01

This table reports the difference-in-difference estimates based on the specification in Equation (3) at the fourth birth order. Column (1) reports estimates for all mothers who give birth to at least four children. Column (2) reports the estimates for mothers who according to our heuristic are expected to sex-select i.e. those whose ideal parity is equal to the given birth order ($n^* = 4$), and columns (3) and (4) do the same for placebo samples. The controls include mother's age at previous birth, rural, wealth quintile, NFHS round-specific fixed effects and state fixed effects. Robust standard errors included in parentheses.

C. CONCEPTUAL FRAMEWORK: FERTILITY PREFERENCES

I model parents to be unitary non-gendered agents who, given their fertility preferences and concurrent birth history, decide whether to conceive again, and whether to sex-select upon conception.[12] Any parent with birth history h could stop childbearing and remain at h or could transit to either $(h, \texttt{F})$ or $(h, \texttt{M})$. Sex-selection is a choice that parents can exercise to limit the possible paths at any history to either remaining at h (by selective abortion) or transiting only to $(h, \texttt{M})$.

Parents begin with a null history or an empty sequence, and based on their actions and birth realizations settle over time into various 'nodes' representing their birth histories—capturing the order and sex of each child born, as well as the sex-composition up to each birth order.

Preferences

Parents' preferences are taken to be a complete and transitive ordering over all possible sex-sequences including an empty sequence representing no births. At the minimum, this ordering is based on parity i.e. the number of elements in the sex-sequence. For several other parents, son-preferring or not, it would also depend on the sex-composition.

Let n^* be the ideal parity for a parent, defined as the total number of children in the most-preferred sex-sequence of the parent. With an assumption that parents monotonically prefer lower parities (all else equal), n^* is well-defined and unique for each parent.[13] Likewise, the 'ideal' number of sons or daughters would be the respective frequencies of M or F in their most preferred sex sequence.[14] Let b^* be the ideal number of sons for the parent, defined as the number of boys in the maximal sex-sequence of the parent's preference ranking.[15]

Examples: For a parent who prefers a son and is otherwise indifferent over parity, based on the definitions, $n^* = b^* = 1$. For a parent who wants equal number of boys and girls, $n^* = 2, b^* = 1$. For a parent who wants two children and is indifferent about the composition, $n^* = 2$ and b^* is undefined.

[12] I abstain from modeling intra-household differences in fertility preferences or bargaining dynamics between spouses (see, for example, Rasul (2008))

[13] If P^* is the set of most preferred or maximal combinations of (n, b), then $n^* = \min\{n | (n, b) \in P^*\}$.

[14] This implicitly assumes that parental preferences are invariant to the particular order in which sons and daughters are born.

[15] For $B = \{b | (n^*, b) \in P^*\}$. I take $b^* \in B$ iff $|B| = 1$, otherwise, b^* is undefined for parents who are indifferent about the sex-composition at n^*.