Nakusha? Son Preference, Resource Concentration and Gender Gaps in Education

Ashwini Deshpande and Apoorva Gupta*

August 10, 2019

Abstract

Using nationally representative data over three decades for India, we investigate mechanisms underlying parental motivation to invest differentially between their sons' and daughters' education. We identify three channels: strong son preference, meta son preference (resulting in "unwanted girls"), and the resource concentration motivation. We find that gender gaps in educational quantity outcomes have declined significantly for all children. However, gender gaps in the quality of education have increased over the period, and the increase is the largest in families with unwanted girls, i.e. due to an intensification of meta son preference, followed by families motivated by resource concentration.

JEL Classification Codes: I21; J16; N35; O15
Keywords: Gender, Education, School Choice, Stream Choice, Educational Expenditure, Discrimination, India

*Deshpande: Department of Economics, Ashoka University, Plot No. 2, Rajiv Gandhi Education City, National Capital Region, P.O.Rai, Sonepat Haryana-131029. India. Email: ashwini.deshpande@ashoka.edu.in. Gupta: Department of Economics, Ramjas College, University of Delhi, Delhi – 110 007. Acknowledgements: The basic idea underlying this paper was explored in one chapter in Gupta's Ph.D. thesis. This paper builds upon the initial work with a substantively reconceptualized analytical framework. Versions of the chapter were presented at Graduate Students' Workshops at the Delhi School of Economics. The paper version was presented at seminars at Indian Institute of Technology, Delhi; CDE (Delhi School of Economics)-CAGE (Warwick) workshop at Delhi School of Economics; Hong Kong University of Science and Technology, and London School of Economics (LSE). We thank Carla Chan, Bhaskar Dutta, Clement Imbert, Naila Kabeer, Ravinder Kaur, Anirban Mukhopadhyay, Albert Park, Rajesh Ramachandran and Sujata Visaria for very useful comments. The usual disclaimer applies.

In rural ares of the state of Maharashtra in India, several parents name their girls "Nakusha" or Unwanted...[1]

1 Introduction

1.1 Son Preference and Resource Concentration

Gender gaps in several health and educational outcomes are ubiquitous, and these are larger in developing countries compared to developed, both due to supply and demand side factors (Jayachandran, 2015). In several countries, a common umbrella, catch-all explanation for demand-side or household-based reasons that account for larger gender gaps in health and education is "son preference". In this paper we examine whether, and to what extent, son preference is a factor in gender-differentiated investments in children's education. We assess the role of son preference relative to other alternative explanations, specifically parental desire to invest more in the child most likely to succeed. We use the newly introduced concept of "unwanted girls" (Jayachandran, 2017) to delineate these explanations, and thus, make a methodological contribution, which enables a more nuanced understanding of the omnibus category of son preference, as well as offers deeper insights into the multiple dimensions characterising gender gaps in education.

The desire for a son, or for a minimum number of sons, and its corollary, daughter aversion, stems from a variety of economic and cultural reasons. The practice of patrilocality implies that the daughter moves into the home of her in-laws, whereas the son stays with the parents after his marriage. Thus, sons are expected to provide old-age support to parents. Daughters will not only move out at the time of marriage, but would need appropriate dowry payments to secure a suitable match. In societies such as India, where marriage is nearly universal, these factors combine to produce strong son preference and daughter aversion. In addition, there are essential religious rituals that only sons can perform.

While a large number of families would desire a minimum number of sons for these reasons, the propensity of parents to discriminate between boys and girls, if they had both, could vary widely. On the one hand, there would be parents with such deep-rooted son preference that they would intrinsically have a negative bias against their daughters, even if they had only girls, and would certainly discriminate against girls, if they had both boys and girls. On the other hand, there would be parents with no "taste for discrimination" against girls, and would intrinsically value equal quality for both sons and daughters.

The most obvious manifestation of strong son preference is sex selective abortions (SSA).

[1] http://www.bbc.com/news/world-south-asia-15414796 (accessed on August 16, 2018)

Thus, families with very strong son preference might already practice SSA to achieve the desired gender balance among their children. However, not all families with son preference practice SSA to proactively alter the sex composition of their children. There would be parents who continue to have children in the hope of getting a son. These families exhibit a "Meta Son Preference". In their desire to keep trying for a son, many of these families might end up with more daughters than they wanted; these would be families with "unwanted girls", following Jayachandran (2017). Such families might exhibit a bias against their "unwanted" daughters, regardless of whether they originally had a taste for discrimination against girls in general.

Finally, there would be families that neither have any specific desire for a minimum number of sons, i.e. no son preference, nor any intrinsic motivation to invest lower amounts in the girl child. However, the absence of son preference might not be sufficient to ensure the absence of gender bias in the investments in the education of their children. The "Resource Dilution Hypothesis" suggests that families are averse to diluting their resources by spreading them evenly across all children, and tend to concentrate their resources on children most likely to succeed. Thus, they might invest less in younger girl children relative to older, or spend less on their daughters' education relative to their sons (Lancaster et al., 2008; Lee 2008; Fors and Lindskog, 2017). In a patriarchal and patrilineal society, the child most likely to succeed would be the son, rather than the daughter, and therefore, such families would invest more in their sons, but not out of a taste for discrimination against girls.

We argue that for a nuanced understanding of the mechanisms underlying gender-differentiated parental investments in their children's education, we should distinguish between a) strong son preference (SP) that manifests itself in absolute preference for boys over girls; and b) lower investments in girls' education compared to boys' for other reasons. While both these tendencies are likely to get exhibited in tandem, the two are not identical. The stronger the former, the more likely are families to actively control the gender composition of their children through the illegal, but widely prevalent, practice of SSA.

There are two methodological challenges to understanding the role of strong SP and other motivations as factors resulting in gender differentiated investments by parents. One, our data do not allow us to comment on the extent of SSA , as we do not have information on the full history of pregnancies and live births. We can only examine any gender bias towards living children, with the caveat that any of the families in our data could have practiced SSA. In other words, we observe the actual number of children and their gender composition at the time of the survey, and can not comment on whether the observed number and gender composition is shaped by SSA or not.

Two, none of the motivations outlined above are directly seen in observational data, which only reveal the contemporaneous family composition. Thus, we can observe three types of

families in the data: those with only girls (GG), those with only boys (BB), and those with both girls and boys (GB or BG). Using a novel methodology explained in Section 2, we divide the latter into two groups: families with unwanted girls, and those without unwanted girls. Based on this classification, we get six types of children (*childtype*): boys in BB; girls in GG; boys in mixed families, and girls in mixed families (with no unwanted girls); boys in families with unwanted girls, and girls in families with unwanted girls. Using data from three special rounds of the National Sample Survey (NSS) focused on education, covering the period 1986-2014, we compare differences in various educational outcomes between the *childtypes* in order to delineate the mechanisms underlying gender gaps, and comment on the extent to which the gaps are due to strong son preference, meta son preference and resource concentration.

1.2 Main Results

We find that over the nearly three decades covered by our study, gender gaps in the quantity of education have either been eliminated or have declined to the point of being insignificant. We estimate three measures for education quantity: probability of ever enrolling in school; of being currently enrolled, and years of education for children between 6 to 19 years. In these indicators, there is a clear and significant decline in gender gaps, across all *childtypes*.

The overall gender gaps in quality of education, as measured by school choice (estimated separately for parents choosing private schools, English-medium schools and private English-medium schools) have increased over the period. This increase is the sharpest for children in families with meta son preference (meta SP), followed by families motivated by resource concentration. The results for educational expenditure (measured through a two-step Heckman selection model, incorporating the estimated probability of current enrolment in the first step) show that conditional on being currently enrolled, gender gaps in educational expenditure were in favour of girls in 1986-87 and 1995-96. However, these have moved sharply against girls between 1995-96 and 2014. This worsening is the highest for unwanted girls, i.e. in families with meta SP, followed by families motivated by resource concentration.

While the relationship between parental motivation and quantity and quality of education is the focus of this paper, we also examine if there is any systematic association between *childtype* and their stream choice in high school. In particular, we estimate the determinants of the gender gaps in the choice of STEM (science, technology, engineering and mathematics) subjects, and how these have evolved over time across the six *childtypes*. We are aware that stream choice by high school students may not be a direct result of parental decisions, and students are old enough to decide based on their aptitudes and factors in the wider school environment, especially the influence of peers. However, international evidence indicates that despite being just as good, women and girls are underrepresented in STEM subjects

(Ceci et al., 2009). This could be the combined reflection of parental attitudes as well as wider societal attitudes or inputs. Our analysis sheds some light on whether parental motivations of son preference and resource concentration also have wider indirect ramifications, for instance, in influencing stream choice.

To the best of our knowledge, ours is the first study that digs deep into understanding the varied motivations behind gendered parental investments in their children's education in the context of a large emerging economy with deep-rooted son preference. We provide evidence on how these motivations have changed over a period of rapid economic growth and widespread structural change. The interpretation of these results in terms of what they signify is more complicated. In Section 7, we turn to the ethnographic literature on the changing nature of son preference over the last few decades in India to understand whether the intensification of meta SP and resource concentration can be understood as a straightforward intensification of gender discrimination, or if these are new strategies towards upward social mobility being adopted by aspirational Indians. We also comment on the changing stereotypes attached to girls and boys as sources of old-age support.

1.3 Context and Contribution to the Literature

India constitutes an ideal setting to examine these issues, as the prevalence of son preference is most obviously manifested in the skewed sex-ratio, which has prompted analyses of "missing women" (Sen, 1990). India also has large gender gaps in educational attainment: Census 2011 shows that the literacy rate among Indian men is 82.14 percent, whereas for women it is 65.46 percent, and 12.8 percent of men have completed secondary level education, compared to 9.6 percent of women.[2] The span period of close to three decades covered by our study coincides with high growth in the Indian economy, as well as by an overall increase in school enrolment for all children. We are, thus, able to comment on the changes in gender gaps in educational attainment in the context of these changes.

Looking at the experience of other countries that managed to reduce manifestations of strong son preference, we see that whereas China (under Mao Tse Tung) and North Korea achieved a decline in son preference through targeted, explicit state-sponsored interventions towards gender equality, the feasibility (and some might argue, desirability) of replicating such interventions in other countries is limited, as the changes were implemented through a control of citizens' private lives. South Korea and Bangladesh experienced a decline in son preference mainly due to structural economic transformation and societal change. As the pre-industrial social organisation in South Korea disintegrated with rapid urbanisation, increasing female education and participation in the labour force, the relationship between parents and their children changed in certain key dimensions. One, daughters were econom-

[2]http://www.mospi.gov.in/publication/women-and-men-india-2014 (accessed on October 18, 2017)

ically as capable to provide parental support as were sons; and two, whether old age care would be provided by the son or the daughter depended more on who lived closer to the parents. Both these factors helped undercut the material basis for son preference (Chung and Das Gupta, 2007). Bangladesh too experienced a similar transformation (Kabeer et al., 2014), discussed in greater detail in Section 7.

A quick comparison with South Korea and Bangladesh reveals that India, in the last three decades, has had in place the requisite setting to enable a decline in son preference. The Indian economy has been undergoing extensive structural transformation over the last three decades, as manifested in greater urbanisation, migration, greater diversity in sources of livelihoods and a movement away from traditional farming occupations. This has been accompanied by various government schemes aimed at enhancing the value of a daughter through subsidies and other monetary incentives, along with changes in inheritance laws. Additionally, in order to influence public perception, there have been vigorous media campaigns emphasising that daughters are just as capable as sons.Yet, India does not seem to have achieved a similar decline in son preference as witnessed in the other two countries. This makes an examination of the Indian context intriguing and necessary.

While there is a great deal of literature on son preference as sex-selection, there is comparatively less literature on understanding broader manifestation of son preference. Our paper advances the latter body of literature by focusing on the effect of "family type" on gender gaps in education. We not only compare educational outcomes separately by family type, ours is the first paper to separately examine the educational outcomes of "unwanted girls". We examine gender gaps in five indicators of education (ever enrolled, currently enrolled, years of education, quality of education proxied by school choice, and household expenditure on education) as well as how the effect of family or child type has changed over time.

Earlier literature on intra-household gender inequality in education in India has examined either household expenditure (Deaton 1989, Subramanian and Deaton 1991, Lancaster et al. 2008, Kingdon 2005), or school choice (Sahoo, 2017). The results of the expenditure studies are discordant, in that some find a pro-male bias and others do not, due to differences in data sources and methodologies. Sahoo (2017), based on panel data for one state in India, finds a gender gap of six percentage points in the probability of private school enrolment among 6-16 year olds, which has been rising over time, providing evidence for greater parental investment in sons' education relative to daughters.

While we explore son preference and alternative explanations in this paper, we should note that various reasons have been attributed to why and how many women are "missing". Son preference and gender discrimination are obviously a part of this story, but the larger body of research shows that there are multiple explanations, such age and disease (Anderson and Ray, 2010). Similarly, as far as gender gaps in education are concerned, parental

attitudes about son preference are an important, but only one part, of the big picture. Deshpande and Nordman (2018) explore the role of neighbourhood and community-level factors as well as gender norms, which would differentially affect girls' ability to continue education, compared to boys. These include distance to school, neighbourhood infrastructure, school-related social networks, exposure to mass media and so forth, after accounting for household-level constraints due to gender norms expressed via expectations about household chores (e.g. time spent by boys and girls on household chores). While in this paper we focus on understanding parental motivation as one of the key factors in explaining gender gaps, it would be important to keep in mind that both demand-side (viz., parental motivation) and supply-side factors would constitute the complete explanation for gender gaps in education.

The rest of this paper is organised as follows: Section 2 presents our methodology for identifying parental motivation. Section 3 describes the data and presents summary statistics. Section 4 presents the estimation results. Section 5 discusses an indirect implication of parental and societal attitudes. Section 6 addresses robustness and selectivity concerns. Section 7 discusses an alternative interpretation of the results. Section 8 offers concluding comments.

2 Methodology

2.1 Identifying Family and Child Types

What we observe in the data are the ex-post number of children, and not the ex-ante desires of parents. How can we infer anything about parental motivation based on observational data?

The literature on son preference identifies "stopping rules", according to which, families with meta son preference keep having children till at least one son (or the desired number of sons) is born. When there is no such rule, sex ratio at birth will be 1.05 whether or not the child is the last one. However, in the presence of this fertility stopping rule, the sex ratio of last child (SRLC) is heavily skewed towards boys, whereas the sex ratio at earlier birth orders would be skewed towards girls. Such meta son preference gives rise to "unwanted" girls: girls born in the process of parents trying for a boy.

We divide "mixed" families into two groups, starting with the birth order of 3, i.e. with families having 3 or more children:

1) If the last child is a boy, and all preceding children are girls, we label such families as "mixed families with unwanted girls". For a family with 3 children, the order of children would be GGB; for n children, $GGG...G_{n-1}B$

2) All other families with 3 or more children having both boys and girls.

The ex-post number and gender-mix of children is a function of desired family size, son preference as well as sheer luck. It is not possible, with observational survey data, to determine with precision whether any particular girl is "wanted" or "unwanted". We have created the category of unwanted girls only for mixed families, which is a lower-bound estimate. For instance, several of the "only girls" families might be families with unwanted girls, where parents kept trying for a boy, never got one, and stopped at some point. Or, parents who first have a girl and then a boy might have wanted both boys, but stop after the boy because they do not want a third child. Thus, that one girl might also be "unwanted", but in our classification will not be captured as such. Further, families with two children might already be the result of SSA, as the concept of small families with one son is increasingly seen as ideal (Basu and Desai, 2016).

Thus, from the observed gender composition of families, we create four family types: GG, BB, Mixed and Unwanted. This implies that there would be six *childypes*: B in BB; G in GG; B_mix; G_mix; B_unw; G_unw.

2.2 Parental Motivation: Family Types

As stated in the Introduction, families would vary in terms of their propensity to discriminate against girls for a variety of reasons. We propose the following schema to understand parental motivation.

Family Type 1 (FT1): Families with 'strong' son preference (Strong SP), either due to a Beckerian "taste for discrimination" against the girl child, or a patriarchal belief in male superiority extending beyond the desire for an optimal number of sons. These families would be explicitly motivated to invest less in their daughters compared to their sons, if they had both sons and daughters. If they have only daughters, such families are less likely to treat their daughters' education as an important and desirable investment, compared to families with the same motivation who have only sons.

Family Type 2 (FT2): Families with meta son preference (Meta SP): parents want at least one son, or more sons than daughters. However, they might not have a "taste for discrimination" against girls, and would value investments in both sons and daughters equally. However, in the process of trying for a son, they might end up having "unwanted" daughters. Thus, they would have a greater than desired family size, with a concomitant squeeze on resources, which would get reflected in lower investments in the unwanted daughters.

Family Type 3 (FT3): Families with a motivation for 'resource concentration': these

would be families with no inherent son preference and no taste for discrimination. In other words, these families would neither have any specific desire for a minimum number of sons, nor any intrinsic motivation to invest lower amounts in the girl child. However, because these families are averse to diluting their resources by spreading them evenly across all children, such families would invest more in their sons, but not out of a taste for discrimination against girls.

Family Type 4 (FT4): Finally, there would be families that have no intrinsic son preference, in the sense of an absolute desire for sons over daughters, and no motivation to treat their sons and daughters unequally.

While the distinction between the four types of families discussed above might be theoretically clear, to establish these differences empirically is challenging. What we observe in the data are *ex-post* number of boys and girls. The data do not reveal how actual family size and sex composition of children are related to intrinsic motivations and preferences of parents, and in particular, if parents have consciously manipulated the sex composition of their children. While prenatal sex determination is illegal in India, it thrives illegally. Section 2.4.2 explains how we associate observed gaps in *childtype* with parental motivation.

2.3 Outcome Variables: Quantity and Quality of School Education

The main objective of this paper is to examine the effect of parental motivation on gender gaps in quantity and quality of school education. We estimate the household-level determinants of the following educational outcome variables for the six groups of children: i) whether the child was ever enrolled in a school (ever enrolled); ii) for 6-16 year old children that ought to be in school, whether the child is currently enrolled;[3] and iii) and years of education to capture educational attainment.

For currently enrolled students, we examine school choice by estimating whether the child goes to a private school[4], English-medium school, or a private English medium school, which are perceived to be of better quality compared to vernacular and/or government-run schools. English is the language of the elite in India, and with increasing globalisation of the Indian economy, English language skills play an even bigger role in securing better outcomes (Munshi and Rosenzweig, 2006). Labour market returns to English language schooling are higher than returns to vernacular schooling (Azam et al., 2013). Private schools are significantly more likely to exist where the public education system has failed to provide good quality

[3]This is a proxy for drop outs, and does not conclusively establish that the child has not discontinued education. This is because the child may be enrolled at the time of the survey, but may have stopped going to school and for all practical purposes, might have discontinued education.

[4]Private schools are privately owned, managed and funded.

education (Kingdon, 2007). Moreover, private schools are, in principle, more accountable to the parents who pay higher fees, and possibly have to exert themselves harder to provide better quality services to the students (Sankar, 2008). English-medium schools and private schools are not synonymous, but their overlap denotes, or is believed to denote, better quality. Desai et al. (2009) show that children who attend private schools and obtain private tutoring are also more proficient in schooling outcomes, i.e. they perform better in tests on reading, writing and arithmetic.Thus, parental choice in sending their children to English medium private schools, which are certainly more expensive than their counterparts, signals their desire to invest more in their children's education, and therefore, their desire for better outcomes for their children.

The NSS has separate data on whether the child's school is English-medium school, as well as on whether the school is private. We show estimates for private school and English-medium separately. Additionally, we club these two variables together to create a binary variable indicating if the child attends private English-medium school or not. Round 42 for 1986-87 did not collect information on medium of instruction, hence the composite quality variable can be constructed for only two rounds: 1995-96 and 2014.

As noted in the Introduction, each of these outcomes would result from a combination of demand-side and supply-side factors; this paper focuses on the demand side factors, and in particular on the role of parental motivation.

2.4 Estimation

2.4.1 Part 1: Overall gender gaps over time

In order to estimate the evolution of gender gaps over time, several of the outcomes variables are dichotomous, and are estimated using a Probit model:

$$P(Y_i = 1|X_i) = \Phi(X_i\beta) \tag{1}$$

where Y refers to the outcome variable (probability of being ever enrolled, currently enrolled, being enrolled in a private school, being enrolled in an English medium school, being enrolled in a private English medium school and stream choice in high school) and $\Phi(.)$ is the standard cumulative normal distribution function.

This is estimated through

$$P(Y_i = 1) = \beta_0 + \beta_1 female + \beta_2 round_k + \beta_3 female * round_k + \beta_4 X_i + \epsilon_i \quad (2)$$

where "round" is the dummy for NSS round; X is the vector of control variables: age, number of siblings, dummy for eldest child; father's and mother's years of education, urban dummy, caste group dummies (Scheduled Caste (SC), Scheduled Tribe (ST), Others (everyone else), real monthly per capita consumption expenditure (MPCE) quartiles, zone dummies.

The coefficient of the interaction term β_3 is the difference-in-differences estimator that shows the change over time (relative to the base year 1986-87) for girls versus boys.

Years of education is a continuous variable, and is estimated using Ordinary Least Squares (OLS) with the same set of explanatory variables as above:

$$EduYears = \beta_0 + \beta_1 female_i + \beta_2 round_k + \beta_3 female * round_k + \beta_4 X_i + \epsilon_i \quad (3)$$

Expenditure on Education

The expenditure on education is a continuous variable, and could be estimated using OLS. However, the dependent variable is observed only for currently enrolled children. To take care of this, we estimate a standard Heckman two-step selection model, described below, where education expenditure is estimated conditional on current enrolment.

Stage 1: Selection Equation, estimating the probability of being currently enrolled, given as follows:

$$P(CurrentEnrol_i = 1|X_i) = \Phi(X_i\alpha') \quad (4)$$

where X_i is the vector of regressors, which include child type, round dummies, age, interaction of child type with round dummies, number of siblings, father's and mother's years of education, urban dummy, caste group dummies (SC, ST, Others), MPCE quartiles, zone dummies, and dummy for eldest child.

Stage 2: Education Expenditure Equation, estimating average education expenditure on a child, conditional on enrolment, given by:

$$EduExp_i = \gamma_0 + \gamma_1 female_i + \gamma_2 round_k + \gamma_3 female * round_k + \gamma_4 Z_i + \epsilon_i \quad (5)$$

where, the dependent variable is the education expenditure on child i. Z_i is the vector of second stage control variables, which in addition to the estimated probability of selection from Equation 7, includes years of education to capture educational attainment.

2.4.2 Part 2: Taste for Discrimination, Meta SP or Resource Concentration?

Equations 2 and 5 estimate the evolution of overall gender gaps (all boys versus all girls) over time. However, our interest is in examining the evolution of gaps between different *childtypes* in order to delineate the mechanisms that underlie the overall change.

We estimate the same equations with dummies for *childtype*:

$$P(Y_i = 1) = \beta_0 + \beta_1 childtype + \beta_2 round_k + \beta_3 childtype * round_k + \beta_4 X_i + \epsilon_i \quad (6)$$

where *childtype* takes 6 values: BB (base); GG; B_mix; G_mix (boys and girls in mixed families with no unwanted girls); B_unw; G_unw (boys and girls in mixed families with unwanted girls)

Years of education is estimated through OLS, analogously to Equation 6.

Comparing these outcomes for the six "child types" allows us to test for alternative mechanisms, in terms of parental motivation, which would lead them to make differential investments between sons and daughters.

1. To test for the presence of strong son preference or taste for discrimination, we compare the outcome of girls in all girls families (GG) with boys in all-boys families (BB), as parents with a strong son preference or strong taste for discrimination would internalise these attitudes to the point of investing less on girls.

2. To test for the presence of Meta SP, we compare outcomes of G_unw with B_unw (girls with boys within families with unwanted girls).

3. To test for the presence of the Resource Concentration mechanism, we compare outcomes of G_mix with B_mix (girls with boys within mixed families).

We should also note another issue related to inference about gender bias in educational expenditure. Since the decision on how much to spend on the education of a child reflects two joint decisions, viz., enrolment and amount, it is possible that gender bias affects the two decisions differently. Specifically, gender bias could result in lower enrolment for girls compared to boys, but conditional on enrolment, families may recognise that girls have special needs in the terms of travel and clothing, and might spend relatively more on girls compared to boys (Kingdon, 2005). Thus, higher amounts of expenditure on girls conditional on enrolment may not necessarily imply the absence of gender bias against them.

3 Data and Descriptive Statistics

3.1 Data

We pool data from three special educational surveys of the National Sample Survey (NSS) for 1986-87 (Round 42), 1995-96 (Round 52) and 2014 (Round 71).[5] In addition to the standard household and individual level data, these rounds have special modules on educational details of individuals between 5 and 29 years[6] and details of expenditure on various heads on education. These are nationally representative surveys, covering about 77,037, 72,883 and 65,926 households in the three rounds respectively.

In NSS data, households, which typically comprise of more than one nuclear family, are primary sampling units (PSUs), and relationships within the household are identified in terms of relationship to the head of household, not in units corresponding to each nuclear family. Since we are focusing on parental motivation, we use a subset of the data where we can match children to their biological parents. Thus, our working sample consists of nuclear families of the household head with 3+ biological children, in the age group of 0-19 years, within which we examine educational outcomes for 6-19 year olds. Section 6.2 discusses the reasons underlying the choice of sample and whether or not it introduces any selection bias in our estimates. While the estimation is done over children in the school-going age, viz., 6-19 years, in order to classify families and children into "childtypes", we take into account the entire birth history of the family. Thus, if a family has three children, only two of which are in the 6-19 age group, that family would be part of our working sample. Similarly, if a

[5]The NSS conducted another special educational survey for the year 2007-08 (Round 64). We decided not to use this round, as this sample seemed like an outlier compared to the three rounds that we did use. For instance, there were a significantly higher proportion of females and female-headed households, which appeared oddly out of line compared to the trend.

[6]Round 71 focused on the age group 5-29; Round 52 on the age group 5-24 and Round 42 did not specify any age.

family has four children: *GGG* in the school going age and one boy younger than six years, the boy would not feature in the estimation, but the family would be classified as "unwanted", rather than as GG.

3.2 Descriptive Statistics

Table 1 shows the composition of the sample for each NSS round by family type and child type. These are weighted estimates reflecting the population that the NSS sample represents.

Our total data covers 80909 families with 3+ children over the three rounds, with a total of 321137 children. As explained above, our working sample covers 225629 children between 6-19 years, whereas the family might have additional children younger than six years. In each round, mixed families are the largest category (around 75 percent of the total); the proportion has declined by roughly 4 percentage points, from 77 to 73 over the period. The proportion of families with unwanted girls has almost doubled from 5.53 to in 1986-87 to 10.52 in 2014. The proportion of GG families has also increased from 3.93 to 5.19 over the period, with a reduction in BB families from 13.93 percent to 11.23. Overall, in our working sample, 78.4 percent of children are born in mixed families, as the bottom panel of Table 1 demonstrates.

Tables 2 and 3 shows the summary statistics, for the key variables used in the analysis, by family type and *childtype* respectively, for the three NSS rounds.

The first panel of Table 2 shows that the average number of children in 3+ families has declined from 4 to 3.7, and this decline is driven by the average number of boys, which has declined from 2.3 to 1.99. Education levels of both parents have increased over the period, with years of education of fathers higher than for mothers (5 versus 2.7 respectively in 2014)[7]. Consistent with the rural-urban distribution of the Indian population, the larger percentage of families with 3+ children across all families is rural (22 percent urban in 2014).

For caste groups, we report averages for Scheduled Castes (SC), Scheduled Tribes (ST), and everyone else ("Others"). It is only in 2014 (Round 71) that the "Others" category was sub-divided into Other Backward Classes (OBC) and the residual "Others", (the latter a proxy for upper castes). Since we do not have the OBC category in two of the three rounds, in the interest of consistency across rounds, we have clubbed OBC and Others into "Others"

[7]NSS collects data on highest level of education attained. Following Deshpande and Ramachandran (2019), we have re-coded educational attainment into years of education. To remain consistent across rounds, we consider the following schema: illiterate or never enrolled = 0 years of schooling; without formal schooling or below primary = 3 years; primary = 5 years; middle = 8 years of schooling; secondary =10 years of schooling; higher secondary = 12 years; graduate = 15 years and post graduate and above = 17 years.

for 2014. The proportion of 3+ families that are SC and ST have increased (from 19 to nearly 22 percent, and 8 to nearly 12 percent respectively) over the period, whereas the proportion of non-SC-ST have declined from 72 to 66 percent over the period. The total proportion of female headed households has increased from 5.8 to 7.2 in 2014.

Income data in developing countries are notoriously hard to collect; we use the monthly per capita expenditure (MPCE) as a proxy for income[8]. The survey contains MPCE data in nominal amounts, which we have converted to real values to make the expenditure values for 1986-87 (NSS 42) and 1995-96 (NSS 52) comparable to 2014 (NSS 71).[9] The average real MPCE has risen over time for each family type to stand at INR 1309 in 2014. We also divide families into MCPE quartiles, and report percentages in each quartile separately. The proportion of 3+ families in the bottom most MPCE quartile have increased (from 27.5 to 41.6 percent), where these have declined for the other three quartiles, indicating that richer families are moving to a norm of two or fewer children, the implications of which we discuss in Section 6.

The next four panels of Table 2 present these summary statistics by the four family types. We see that the proportion of urban GG families is higher than average (nearly 25 percent in 2014), as well as proportion in the lowest MPCE quartile (44.6 percent in 2014), whereas the average MPCE in 2014 is not significantly different than average. Mothers in GG families are better educated than average (3.8 years in 2014), wheres fathers' education years is similar to the total average. The highest proportion of GG families is in the Eastern zone.[10]

The last panel of Table 2 presents the summary statistics for mixed families with unwanted girls. We see that these families have better educated parents, with fathers' average years of education at 6.3 and mothers' at 4.2 in 2014. These families have a higher than average proportion of urban families (27 percent in 2014), as well as higher average MPCE

[8]The special education rounds of the NSS, which we have used in the analysis, do not contain a great deal of data on household assets. The survey collected information on land ownership in 1986-87 but not in subsequent rounds, thus we cannot assess change over time. Results (available upon request from authors) show that mixed families (with no unwanted girls) owned less land than all-boys families. Similarly, information on the ownership of a computer is collected only for 2014, and not in earlier rounds. Our tabulations (not reported) show that the overwhelming proportion of families do not own computers. Across family types, GG and Mixed families with unwanted girls form the largest proportion of families that own a computer (roughly 10 percent and 11.2 percent respectively)

[9]We use Consumer Price Index for agricultural workers (CPI-AL) for rural families, and Consumer Price Index for industrial workers (CPI-IW) for urban families to construct the constant price series.

[10]We divide states into six geographical zones: North (Delhi, Haryana, Jammu and Kashmir, Himachal Pradesh, Uttar Pradesh, Uttarakhand, Punjab); South (Andaman and Nicobar Islands, Andhra Pradesh, Karnataka, Kerala, Lakshadweep, Pondicherry, Tamil Nadu); East (Bihar, Jharkhand, Orissa, West Bengal); West (Dadar and Nagar Haveli, Daman and Diu, Goa, Gujarat, Maharashtra, Rajasthan); Central (Madhya Pradesh and Chhattisgarh); and North- East (Sikkim, Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Tripura, Nagaland).

(INR 1435 in 2014). All these features suggest that meta son preference is not a phenomenon confined to poorer families, and is consistent with the evidence that strong son preference that results in SSA is more prevalent in richer, urban families.

Information on fathers' occupation (another proxy for socio-economic differentiation) was collected in two of the three rounds (1986-87 and 2014). On average, there has not been significant change in this variable, although the distribution across family type shows significant differences, both over time and across family type. Table 2 shows that a higher proportion of fathers in families with unwanted girls are employed in high-end occupations like legislators, professionals, or technicians compared to fathers of other family types (GG, BB, or Mix), and this proportion has increased over time, from 8.3 percent in 1986-87 to 14.51 percent in 2014. Moreover, a comparatively lower proportion of fathers in mixed families are employed in these high-end occupations (6 percent in 1986-87, and 10.5 percent in 2014).

Table 3 presents the summary statistics for the six *childtypes*, the differences in which we analyse in the next section.

4 Results

This section presents the results for the estimation for the various educational outcomes.

4.1 Quantity of Education

4.1.1 Ever Enrolled and Currently Enrolled

Figures 1 and 2 show the contrast, between girls and boys, in the predicted probability of being ever enrolled and being currently enrolled over the three rounds. These are results from the estimation of Equation 2. These are conditional estimates, with all the controls mentioned in Section 2. We see that in 1986-87, there was a gender gap of 0.16 and 0.09 in the probability of ever enrolling and being currently enrolled respectively. There is a consistent and monotonic decline in the gap between girls and boys in both these outcomes, such that by 2014, we see an elimination of the gender gap in current enrolment.

A pairwise comparison of outcomes by *childtype* can be seen in Figures 3 and 4, which present the conditional estimates from Equation 4. As outlined in Section 2, a comparison of BB and GG in Figure 3 reveals that the gap in the predicted probability of ever enrolment has fallen over the period, with gap being close to zero in 2014. We see a similar pattern for the gap between B_mix and G_mix, as well as between B_unw and G_unw. Thus, all three mechanisms (taste for discrimination, resource concentration and meta son preference respectively) show a clear weakening over the period. In other words, the aggregate results of a decline in the gender gap is a result of the decline in the gender gap across all three

types of families.

Figure 4 reveals a similar picture as Figure 3, with gap between BB and GG being slightly negative (girls more likely to be currently enrolled), and gaps between B_mix and G_mix, as well as between B_unw and G_unw being zero in 2014.

4.1.2 Years of Education

Figure 5 shows the evolution of gender gaps in years of education. In 1986-87, girls' average years of education was roughly 0.65 years less compared to boys. As we would expect based on trends in current enrolment, there is a clear decline in the gender gaps in years of education, such that these are eliminated by 2014. Figure 6 reveals that this decline is seen across all *childtypes*, with the decline in families with unwanted girls being the largest. Thus, we see that in all three indicators of quantity of education, there is a clear convergence between all boys and girls, i.e. clear reduction of gender gaps to the point where gender gaps have been eliminated by 2014.

4.2 Quality of Education

This section presents results on evolution of gender gaps in quality of education, as measured by enrolment in private schools, in English medium schools (and in private English medium schools) both between all boys and all girls, and across *childtypes*.

Figure 7 shows the evolution of gender gaps in private schooling. We see that the probability of girls being enrolled in private schools was 0.02 points lower than that for boys in 1986-87. Between 1986-87 and 1995-96, this gap did not change. However, by 2014, the probability of girls being sent to private schools was 0.05 less than that of boys, in other words, the gap between girls and boys has increased over the period. The straightforward interpretation of this is an increase in gender discrimination; however in Section 7.1, we discuss an alternative, more favourable, interpretation of this gap. Data on English-medium schools were collected only in 1995-96 and 2014, and we see a similar increase in gaps in the probability of girls being enrolled in English-medium schools in Figure 8.[11]

Figure 9, which presents the results for private schooling by *childtypes*, allows us to see which groups of children contribute to this overall trend of increasing gender gaps. We see that the largest increase in gaps is within children of "unwanted" families, i.e. between "unwanted" girls and their brothers. We see the same trend in Figure 10, which depicts

[11] Combining the two indicators (private schools and English medium schools), we created a composite indicator of private English medium schools, and that shows the same trend. These results (not reported) are available with the authors upon request.

the results for the conditional probability of going to English-medium schools by *childtype*. Thus, the increase in gender gaps in education quality indicators is largely due to an intensification of meta son preference. Putting it differently, these results show that over time meta son preference has adversely affected the quality of education of "unwanted" girls the most. We see an increase in gender gaps within "mixed" families as well, but the increase is much smaller in magnitude.

4.3 Other Covariates

Table 4 shows the average marginal effects of all the covariates for each of the outcome variables discussed in this section. We see that education of both parents increases the probability of better outcomes for girls, relative to boys. Interestingly, the number of siblings negatively affects the quantity outcomes of children, but not the quality variables. Urban children do better than rural on all the indicators. SC children have worse quality outcomes than "Other" children, but their probability of enrolment is not different. They have slightly lower years of education compared to Other children. ST children are worse-off than Other children in all outcomes, except current enrolment and private English medium schooling. Relative to the lowest MPCE quartile, children in the other three quartiles do better on all indicators, with the marginal effect rising monotonically. Children from the South and North-East zones have better outcomes than those in the North, whereas those in the West have worse outcomes and those in the East have mixed.

Table 5 presents the full set of results by *childtype*. The base category is boys in all-boys. The effect of other covariates remains similar as that in Table 4.

4.4 Expenditure on Education

Figure 11 shows the trend in the gender gap in annual educational expenditure (in real terms)[12], which are the results from the second stage of the Heckman estimation, i.e. average amount spent on education, conditional on current enrolment. We see that in 1986-87 and 1995-96, the average expenditure on girls (conditional on being currently enrolled) was INR 256.7 and INR 188 higher respectively, than that for boys. By 2014, female disadvantage had worsened sharply such that parents were spending INR 509 less on girls than on boys. Table 6 presents the full results.

Figure 12 presents the estimates with pairwise comparison of *childtypes*. We see that the gap in educational expenditure, conditional on being currently enrolled, is the highest

[12]We use Consumer Price Index for agricultural workers (CPI-AL) for rural and Consumer Price Index for industrial workers (CPI-IW) for urban families to convert nominal expenditure to 2014 prices.

in "unwanted" families, with parents spending roughly INR 2200 per year more on boys compared to unwanted girls in 2014. Notice that in 1985 and 1995-96, the gender gap was close to zero. In "mixed" families too, we see an increase in gender gaps of roughly INR 600 per year but the increase is of a smaller magnitude. Thus, the largest increase in disadvantage in terms of educational expenditure is due to meta son preference followed by resource concentration. Breaking this down by items of expenditure, we note (not reported in paper, available with authors) that these gaps are almost entirely due to the gaps in private tutoring.

5 Parental and Societal Attitudes: Stream Choice

The schooling system in India requires students to choose one of three broad streams – Humanities, Commerce and Science[13] in Class XI, i.e. for the last two years of high school. The gendered patterns in choice of subjects is seen in several countries across very diverse cultural contexts, such that girls are less likely to choose STEM subjects. We recognize that stream choice for high schoolers may not directly be dictated by parents; however, parental and societal attitudes towards gender discrimination and girls' actual experience of gender bias is likely to shape their subject choices in high school.

Figure 13 shows that in 1986-87, girls were significantly more likely to choose Humanities, and less likely to choose Commerce. Interestingly, girls were as likely to choose Science as the boys. By 1995-96, this pattern changed, in that the gender gap for Humanities went down (but girls remained more likely), and the gender gap in Commerce became insignificant. Within this decade, a gender gap *appeared* in the probability of choosing Science, in that girls were significantly less likely than boys to choose Science. By 2014, this pattern continued, with a further decline in the gender gap for humanities and a slight but insignificant narrowing of the gap for Science. For commerce, the gender gap in predicted probabilities stands at zero.

Figure 14 breaks this pattern up by *childtype*. The increase in the gender gap in Science between 1986-87 and 1995-96 is entirely due to the "unwanted" families, in which boys are more likely to opt for Science compared to unwanted girls. Again, not to put too fine a point on it, i.e. this is not causal by any means, but an increase in meta son preference clearly has other indirect ramifications; specifically, it contributes to girls moving away from STEM subjects in high school.

[13]Actually, the Science stream is further subdivided into Science with Biology, for those potentially interested in studying medicine, and Science without Biology, and with Computer Science or Mechanical Drawing, for those interested in engineering and allied topics. However, the NSS data clubs all Science students in one category.

6 Robustness and Selectivity Concerns

6.1 Possible Selection Bias?

Our working sample consists of biological children of the household head, and within that, we are focusing on families with 3 or more children. Thus, we are excluding children whose parents we cannot identify in NSS data, as well as families with 1-2 children from our analysis. What could be the possible bias from this selection?

India's total fertility rate (TFR) declined from 3.16 to 2.66 between 2001 and 2011 (figures from decennial censuses from Basu and Desai, 2016, p. 27). Basu and Desai (2016) present evidence that in nine out of 35 Indian states, TFR is below replacement level, i.e. less than 2, with a small proportion displaying a trend towards very small families, i.e. with one child. Using data from India Human Development Survey (IHDS) for 2004-05, they estimate that nationally about 5 percent of families have one child out of choice, and about 40 percent of these families seem to have stopped at one child despite this child being a daughter. In other words, they appear to be content with having an only daughter, and do not manipulate the sex of the only child.

One child families appear to be highly motivated towards investments in their children. Estimates from Basu and Desai (2016) reveal that educational expenditure is 40 percent higher in such families compared to families with 3+ children. Children in one-child families are 1.56 times more likely to attend private school compared to children from 3+ families. Interacting sex of the child with number of children, they find that among one-child families, parents do not distinguish between boys and girls. Thus, there is no gender bias in educational investments in one-child families. Overall, these single children are atypical, but highly advantaged in terms of being sent to private schools, to English medium schools, more likely to be aided by private tuitions.

What about 2-child families? In India, this has become the norm over the last two decades (Kaur and Vasudev, 2019). One set of such families might genuinely want one boy and one girl (or may not care about the sex of the two children); in which case, the problem of gender bias would either be absent or minimised. However, in countries like India, fertility has declined faster than son preference, and this has resulted in the "intensification effect" or the increased elimination of girl children at lower order births (Das Gupta and Bhat, 1997), as parents aim to achieve their desired sex composition of children within a smaller family size. Kaur et al. (2016) report studies showing a sharp decline in the girl-to-boy sex ratio for second order births when the first born was a girl (p.9), which indicates SSA at the second birth order. John (2018) documents how son preference (and daughter aversion) pose a challenge, where the small family of one boy and one girl is the norm. She argues that this seemingly egalitarian practice hides more muted forms of son preference, as families are

averse to having both girls if they want two children, and are likely to sex-select.

The implications of this evidence for our study are that within families with less than three children, either parents are completely egalitarian, or that in two-child families, they might have already exercised SSA to achieve the desired gender balance, and are, therefore, not likely to discriminate further against their daughters. Thus, families of 3+ children is the correct subset of families to consider, if the intention is to evaluate the role of parental motivation resulting in gender bias or female disadvantage within living children.

6.2 Why Would There Be "Unwanted" Girls with Widespread SSA?

A question that arises in the context of widespread SSA is this: why would families have unwanted girls at all when SSA is so widespread, and apparently quite easy? Despite the widespread and ubiquitous practice of illegal and underground SSA, not all parents actually practice it. Qualitative interviews reveal that parents often find themselves "resigned to the fate that god has dealt to them" (John, 2018, p. xii). John's study also indicates that due to slightly better regulation, the practice might have become more expensive and clandestine in parts. Kaur and Vasudev (2019) report on the stigma on abortions, as they believe it is going against nature or against the will of God. Thus, as long as families have a meta son preference, and do not practice SSA, there would be "unwanted" girls, with resulting implications for material outcomes, such as educational outcomes that this paper has documented.

6.2.1 Birth Spacing

As we noted earlier, in the absence of the full pregnancy and birth histories of women, we cannot infer anything definite about whether families exercised SSA to alter the size and gender mix of their children.

However, one interesting indicator is birth spacing, which is a rough indicator of parental decisions, as contraceptive use in India has increased over 1992-93 and 2015-16 (Rai, 2017), with the caveat that changes in birth spacing over time could reflect a variety of factors other than son preference. For instance, if both parents are working, they might opt to have their children spaced close, so that the childbearing and childrearing phase, which is likely to disturb their work schedules, would be taken care off in a shorter span of time, rather than extend it over several years. In terms of the relationship between birth spacing and son preference, one hypothesis is that if gap between the first and second child is higher when the first child is a boy, compared to when it is a girl, it is indicative of the resource concentration mechanism. The desire for another child after a boy indicates that parents are working towards a desired family size, but would like to delay the next child in order to provide proper nurture for the boy child. However, if the birth of a girl is followed by the

birth of a boy after a longer than average gap, it is taken as indicative of SSA (i.e. suggests that there might have been a terminated pregnancy in between).

Table 7 shows summary statistics for birth spacing for the four family types across the three rounds at each birth order till the sixth. We see that in 1986-87, in mixed families, when the first child is a boy, the average birth interval between the first and the second child is 4.37 years, whereas it is 3.04 years when it is a girl. In mixed families with unwanted girls, it is 3.21 at the first birth order in 1986-87. By 2014, in mixed families with unwanted girls, when the first child is a girl, the next child is born after 2.87 years, the second one after 3.05 years, the third one after 2.88 years and so on. Between the fifth and sixth, this interval increases to 3.49, which might be indicative of SSA, in that it is possible that there were pregnancies that were aborted.

In mixed families, the birth interval after a boy is definitely longer than after a girl, in each round and at each birth order. This, *prima facie*, appears to be supportive of the resource concentration motivation.

6.3 Stability of Family Types

Since we are using contemporaneous data, we are capturing the birth history of families up until that point. How can we be sure that family types (and hence "childtypes") will not change as parents have more children? How do we know that the last recorded child is, in fact, the last child? Since the data is not longitudinal, we cannot rule out entirely the possibility of additional children in the future. This possibility would affect the classification of some families: specifically, some GG families could end up becoming "unwanted" in the future if parents have a boy and stop; similarly, some BB families could end up becoming "mixed" if the parents have a girl at some point in the future. In principle, some "unwanted" families could end up becoming "mixed" if they have another girl after their last boy, maybe in the hope of another boy.

In order to rule out this possibility, we estimated our results for a subset of families where the youngest child is six years or older.[14] Given the average birth spacing of two years, we believe that the probability of additional children, when the youngest child is six or older, is very small. This lowered the total sample from 225,629 to 143,686 children, with the following distribution of family types (overall, across the three rounds): 3.14 percent GG,

[14]If we had retrospective data, we could have narrowed our sample to mothers old enough to be at the end of their reproductive ages, and used educational data on their children. However, with NSS data, this is not possible, as educational data such as expenditure are collected for children currently enrolled. Given the early age of marriage and childbirth soon after marriage, proportion of mothers at the end of their reproductive age, say 45 years and older, with children under 19 is very small, and focusing on them would have not only have seriously truncated our sample size, but also introduced additional selectivity concerns.

16.01 percent BB; 78.83 "mixed" and 7.02 percent "unwanted". All our results are robust to the use of the smaller sample; estimates remain significant with larger standard errors.[15]

One final point relates to how our family types correspond to explicitly stated son preference, and whether the differences we are measuring do, in fact, assess parental motivation. NSS data do not have indicators on explicit son preference. We this issue briefly in Appendix A, based on our estimates from the National Family and Health Survey data.

7 Discussion: Strategies for Social Mobility?

We started the paper by referring to the name *Nakusha* in the state of Maharashtra in India. The phenomenon of naming girls as "unwanted" is by no means uniquely Maharashtrian. In another state of India, West Bengal, girls are named *An-na*, which is a colloquial version of *Aar Na*, which in Bengali means "no more". In Chinese, girls are named *Dai di*, where the first character means "bring along" and the second one means "younger brother". Lately, parents who name their daughters *Dai di* are often scorned, as it is seen as derogatory towards the girls. Girls are often named *Mei Di*, where the first character means beautiful, but sounds in local Cantonese dialects as "do not have", indicating the disappointment of the parents.[16]

Being "unwanted" has implications that go beyond naming; there are clear consequences for quality of girls' education, relative to their brothers'. We document clear evidence of persistent and increasing discrimination against girls in terms of quality of education, which has been the highest in families with meta son preference (i.e. against unwanted daughters), followed by families motivated by the resource concentration mechanism. Prima facie, this indicates that the combination of patrilocality and near universality of marriage continue to sustain the notion that girls are, what is called *paraya dhan*, in Hindi, meaning literally another's property. Viewed this way, families might perceive that they will not benefit from investment in their daughters' education, as this will be akin to "watering the neighbour's garden" (Attané and Guilmoto, 2007). The deep-rooted and persistent pressure of generating a dowry for girls might also contribute to the notion that investing in higher quality education for girls is a waste of precious resources. Kaur and Vasudev (2019)'s recent qualitative account confirms the role of dowry in daughter aversion. In their field site of two villages in the north Indian region of Jammu, they find that higher dowries are being demanded and given, compared to the previous generation.

[15]These results are available upon request.

[16]The details about the Chinese names are from personal communication with Carla Chan of HKUST.

7.1 Winds of Change

However, there are several other changes happening simultaneously that question a straightforward interpretation of these results. Kaur et al. (2016) argue that shaping the size and sex composition of families could be a part of the social mobility strategies of the new elite, who want smaller families and have aspirations for upward mobility through better performance of their children. The same argument could be extended to gender differentiated investments that are motivated by resource concentration. Additionally, there has been some change in son preference attitudes in India, as witnessed through an improvement in the sex ratio at birth from a peak of 113.6 in 2004 to 110 in 2012, which is still above the natural average of 105, but is an improvement nevertheless. This is noteworthy, despite the fact that the change is slow and uneven.

Qualitative studies reveal the beginning of the emergence of a new gender stereotype: that of a caring daughter, contrasted with sons who cannot be relied upon, especially after their marriage. John (2008), in study of selected districts in five Indian states with highly adverse SRBs, finds parents expressing concern about their sons, who they saw as "worthless" or having fallen into bad habits. Reductions in the distance between parental and marital homes have also assisted the feeling that daughters can just as easily provide critical old age support. There are also accounts of parents desiring a son not so much for themselves, but for the sake of their daughters when they themselves are gone, i.e. daughters must have brothers, as dealing with the marital family would become easier if they had brothers to watch out for them (John, 2018).

Also, there are several schemes aimed at enhancing girls' education, but they are restricted to girls in government schools, such as the bicycle scheme in Bihar (Muralidharan and Prakash, 2017), or the bicycle scheme in Karnataka (Avinash, 2017), or various conditional cash transfers (Sekher, 2012). Thus, the fact that in mixed families, girls get sent to public schools with greater probability could simply reflect parental strategies for optimal use of resources.

7.2 The Bangladesh Story

There is an interesting and instructive parallel to be drawn with India's neighbour, Bangladesh, which has also had a history of son preference and SSA. The early literature on Bangladesh noted the great deal of affinity that Bangladesh had with the north Indian family and kinship system of patriliny, patrilocality, caste endogamy and village exogamy. As Kabeer et al. (2014) note, "women's dependence on men for protection and provision at different stages of their lives gave them a particularly strong stake in producing sons – both to ensure their own place within their husband's kinship group and as a form of security for their old age."

Fertility started to decline in Bangladesh in the 1970s, and this was accompanied by a decline in SSA, unlike in India, where TFR is declining but not SSA, as son preference continues to prevail. Qualitative accounts from Bangladesh (Kabeer et al., 2014, pp. 152-155) indicate the changing gender stereotypes, which have become increasingly favourable to girls, as boys are seen to fall into bad habits, take drugs, escape parental discipline, whereas girls are seen to be capable of providing emotional support, and are considered more empathetic and compassionate. The higher and rising presence of women in the labour force could be one of the factors that explains why Bangladesh has managed to achieve a turnaround in SSA and a marked reduction in son preference, whereas India has not. The normal or more conventional route of women being valued through their economic contribution to the family is not on the cards for India, as at the time of writing, the country is facing a decline in the already low labour force participation rates.

8 Concluding Comments

In societies with widespread and deep-rooted son preference, such as India and China, several girls are born in the hope of a boy. These girls are "unwanted". Using pooled data over 1986-87 and 2014 from the special education surveys of the NSS, we investigate the role of parental motivation in the gendered pattern of investments in their children's education. We investigate if gender gaps in quantity and quality of education can be explained by parental motivation; in particular, if gender gaps depend on whether they have "unwanted" girls. We suggest a methodology to distinguish between channels that might drive parental motivation: strong son preference, meta son preference (resulting in unwanted girls) and resource concentration.

Consistent with other studies (e.g. Kishor and Gupta, 2009), we find that over the last two decades overall, the quantity of education has increased for all children as witnessed by greater enrolment and higher years of education. This increase is seen among girls and boys, across all family types, including in families with unwanted girls. The fact that during this period, successive national and state governments gave a big push towards universalising enrolment through schemes such as *Sarva Shiksha Abhiyan* and the 1986 New Education Policy must have provided a conducive environment for this change to have taken place. The less charitable explanation behind rising female education is that a certain level of education, upto a graduate degree, is seen as the minimum qualification for marriages in families seeking upward mobility (John, 2018).

As the Bangladesh and South Korean stories illustrate, with societal change, and daughters proving their worth in terms of old age support, contribution to the natal family in terms of both pecuniary and emotional support, it is possible for parental motivations to change. However, India, which also moved towards liberalisation of the economy with concomitant

structural shifts around the same time, does not appear to see a similar reduction in son preference that these two countries have witnessed. As we argue above, this could be because of the low economic value of women due to their stubbornly low and declining labour force participation rates.

The results of our study point towards the urgent need for mainstreaming gender concerns in educational and labour market policies, as well as highlight the multitude of positive contributions by Indian women in multiple fields to dispel the notion that investment in the quality of daughters' education is a waste of resources.

Tables

Table 1: The Sample

	Number of families with 3+ children(%, N)				
	(1)	(2)	(3)	(4)	(5)
	Only Girls	***Only Boys***	***Mixed***	***Unwanted***	***Total***
1986-87					
%	3.93	13.93	76.61	5.53	100
N	1172	4882	27242	1700	34996
1995-96					
%	4.35	12.84	75.94	6.87	100
N	1189	3755	21739	1913	28596
2014					
%	5.19	11.23	73.06	10.52	100
N	884	1937	12590	1906	17317
Total					
%	4.5	12.64	75.18	7.68	100
N	3245	10574	61571	5519	80909
	Total Number of children, by family type (Mean, N)				
1986-87					
Mean	3.54	3.54	4.20	3.53	4.05
N	4199	17546	117931	5987	145663
1995-96					
Mean	3.51	3.47	4.00	3.42	3.87
N	4155	13062	88145	6541	111903
2014					
Mean	3.40	3.34	3.82	3.38	3.70
N	2968	6406	47814	6383	63571
Total					
Mean	3.47	3.46	4.01	3.43	3.87
N	11322	37014	253890	18911	321137

	Working Sample, Children between 6-19 years (Mean, N, %)				
	(1)	(2)	(3)	(4)	(5)
	Only Girls	***Only Boys***	***Mixed***	***Unwanted***	***Total***
1986-87					
Mean	2.26	2.16	2.83	2.43	2.69
N	2824	11155	82383	4310	100672
%	3.3	11.2	80.5	5	100
1995-96					
Mean	2.28	2.27	2.76	2.46	2.66
N	2867	8750	63265	4965	79847
%	3.74	10.97	78.94	6.36	100
2014					
Mean	2.41	2.31	2.81	2.53	2.70
N	2093	4313	33979	4725	45110
%	4.63	9.61	75.9	9.86	100
Total					
Mean	2.33	2.25	2.80	2.49	2.68
N	7784	24218	179627	14000	225629
%	3.9	10.58	78.4	7.12	100

Source: Authors' calculations based on NSS data

Table 2: Family Level Descriptives

Total Sample

	(1)	(2)	(3)
	1986-87 (Round 42)	***1995-96 (Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std.Dev.	Std.Dev.	Std.Dev.
No. of children	4.047	3.873	3.700
	1.183	1.082	1.008
No. of girls	1.721	1.716	1.705
	1.168	1.135	1.098
No. of boys	2.326	2.157	1.995
	1.199	1.125	1.064
Father Edu Years	3.662	3.988	5.017
	4.155	4.384	4.697
Mother Edu Years	1.555	1.733	2.673
	2.967	3.198	3.833
Urban	.2264	.231	.221
	.4185	.4215	.4147
SC	.1900	.2285	.2187
	.3923	.4199	.4133
ST	.0827	.0880	.1195
	.2755	.2833	.3244
Others	.7273	.6835	.6618
	.4454	.4651	.4731
HH size	6.860	6.425	6.106
	2.190	1.900	1.594
Female Headed HH	.0585	.0603	.0725
	.2347	.2381	.2593
MPCE (INR)	928.33	1110.20	1308.59
	361.96	450.90	624.85
MPCE_Q1 (%)	.2752	.2828	.4160
	.4466	.4503	.4929
MPCE_Q2 (%)	.2556	.2630	.2357
	.4362	.4402	.4244
MPCE_Q3 (%)	.2506	.2404	.2065
	.4333	.4274	.4048
MPCE_Q4 (%)	.2187	.2138	.1417
	.4133	.4100	.3488
Father's Occupation			
Legislators	.0140	.	.0615
	.1177	.	.2402
Professional	.0183	.	.0226
	.1341	.	.1485
Technicians	.0289	.	.0230
	.1676	.	.1500
Clerks	.0228	.	.0146
	.1494	.	.1198
Services	.0872	.	.0682

	.2821	.	.2522
Agriculture	.3846	.	.3284
	.4865	.	.4696
Craft Workers	.0893	.	.1456
	.2852	.	.3527
Plant Operators	.0400	.	.0562
	.1959	.	.2304
Elementary Occupations	.3147	.	.2800
	.4644	.	.4490
Zone			
North	.2551	.2554	.3211
	.4359	.4361	.4669
South	.2253	.2087	.1151
	.4178	.4064	.3192
East	.2209	.2193	.2517
	.4148	.4138	.4340
West	.1868	.1877	.1726
	.3898	.3905	.3779
Central	.0765	.0873	.1019
	.2656	.2822	.3025
NE	.0354	.0416	.03760
	.1849	.1996	.1902
N	34996	28596	17317

All-Girls Families (GG)

	(1)	(2)	(3)
	1986-87 (Round 42)	***1995-96 (Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std.Dev.	Std.Dev.	Std.Dev.
No. of girls	3.541	3.509	3.397
	.8156	.7693	.7697
Father Edu Years	4.461	4.421	5.441
	4.537	4.666	4.662
Mothers Edu Years	2.279	2.391	3.798
	3.671	3.910	4.200
Urban	.2451	.2715	.2480
	.4303	.4449	.4321
SC	.1865	.2246	.2301
	.3897	.4175	.4211
ST	.1052	.1081	.1099
	.3070	.3107	.3130
Others	.7083	.6672	.6600
	.4547	.4714	.4740
HH size	5.972	5.836	5.660
	1.667	1.462	1.351
Female Headed HH	.0904	.0574	.0759
	.2868	.2327	.2650
MPCE (INR)	952.99	1104.04	1304.77
	398.98	460.78	624.21
MPCE_Q1 (%)	.2915	.3103	.4464
	.4547	.4628	.4974
MPCE_Q2 (%)	.2253	.2516	.1743
	.4180	.4342	.3796
MPCE_Q3 (%)	.2377	.2289	.2318
	.4259	.4203	.4222
MPCE_Q4 (%)	.2455	.2092	.1476
	.4306	.4069	.3549
Father's Occupation			
Legislators	.0213	.	.0636
	.14432	.	.2441
Professionals	.0344	.	.0165
	.1824	.	.1276
Technicians	.0282	.	.0201
	.1657	.	.1403
Clerks	.0245	.	.0112
	.1547	.	.1051
Services	.0937	.	.0914
	.2916	.	.2884
Agriculture	.3614	.	.2701
	.4806	.	.4443
Craft works	.0879	.	.1612
	.2833	.	.3680

Plant operators	.0376	.	.0567
	.1904	.	.2315
Elementary Occupations	.3109	.	.3092
	.4631	.	.4624
Zone			
North	.2105	.1934	.2506
	.4078	.3951	.4336
South	.2799	.2498	.2138
	.4491	.4331	.4102
East	.1953	.2200	.2904
	.3966	.4144	.4542
West	.2137	.2006	.1102
	.4101	.4007	.3133
Central	.0714	.1069	.0935
	.2576	.3092	.2912
NE	.0292	.0292	.0416
	.1683	.1684	.1997
N	1172	1189	884

All-Boys Families (BB)

	(1)	(2)	(3)
	1986-87 (Round 42)	***1995-96 (Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std.Dev.	Std.Dev.	Std.Dev.
No. of boys	3.545	3.465	3.340
	.8243	.7709	.6750
Father Edu Years	3.128	3.711	4.619
	3.949	4.175	4.550
Mother Edu Years	1.172	1.407	2.170
	2.653	2.859	3.475
Urban	.1966	.2146	.2211
	.3975	.4106	.4151
SC	.1917	.2610	.2519
	.3937	.4393	.4342
ST	.0867	.0727	.1218
	.2814	.2596	.3272
Others	.7216	.6663	.6262
	.4483	.4716	.4839
HH size	7.089	6.488	6.143
	2.628	2.267	1.898
Female Headed HH	.0543	.0641	.0748
	.2266	.2449	.2631
MPCE (INR)	934.33	1131.86	1318.43
	352.37	452.59	599.38
MPCE_Q1 (%)	.2565	.2533	.3773
	.4368	.4350	.4849
MPCE_Q2 (%)	.2622	.2688	.2532
	.4399	.4434	.4350
MPCE_Q3 (%)	.2647	.2348	.2273
	.4412	.4240	.4192
MPCE_Q4 (%)	.2165	.2431	.1421
	.4119	.4290	.3493
Father's Occupation			
Legislators	.0156	.	.0572
	.1240	.	.2323
Professionals	.0141	.	.0133
	.1178	.	.1146
Technicians	.0207	.	.0193
	.1424	.	.1377
Clerks	.0191	.	.0119
	.1368	.	.1087
Services	.0845	.	.0788
	.2782	.	.2695
Agriculture	.4281	.	.3361
	.4949	.	.4725
Craft Workers	.0823	.	.1467
	.2749	.	.3539

Plant Operators	.0314	.	.0625
	.1745	.	.2422
Elementary Occupations	.3041	.	.2741
	.4601	.	.4462
Zone			
North	.2779	.2594	.3020
	.4480	.4383	.4593
South	.1940	.1970	.1222
	.3955	.3978	.3276
East	.2260	.2437	.2583
	.4183	.4294	.4378
West	.1923	.1826	.1949
	.3941	.3864	.3963
Central	.0864	.0819	.0883
	.2809	.2743	.2838
NE	.0234	.0354	.0342
	.1511	.1848	.1819
N	4882	3755	1937

Mixed Families (with no unwanted girls)

	(1)	(2)	(3)
	1986-87 (Round 42)	***1995-96 (Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std.Dev.	Std.Dev.	Std.Dev.
No. of children	4.202	4.004	3.823
	1.233	1.136	1.071
No. of girls	1.882	1.840	1.750
	.9570	.9357	.9021
No. of boys	2.320	2.164	2.074
	1.038	.9434	.8767
Father Edu Years	3.661	3.883	4.861
	4.122	4.344	4.683
Mother Edu Years	1.523	1.640	2.439
	2.903	3.079	3.689
Urban	.2296	.2271	.2113
	.4206	.4189	.4082
SC	.1909	.2283	.2163
	.3931	.4198	.4117
ST	.0812	.0910	.1219
	.2731	.2876	.3272
Others	.7279	.6807	.6619
	.4451	.4662	.4731
HH size	6.934	6.511	6.200
	2.142	1.878	1.597
Female Headed HH	.0572	.0603	.0720
	.2322	.2380	.2585
MPCE (INR)	923.01	1099.82	1288.87
	358.23	444.45	611.20
MPCE_Q1 (%)	.2778	.2912	.4290
	.4479	.4543	.4950
MPCE_Q2 (%)	.2571	.2619	.2402
	.4370	.4397	.4272
MPCE_Q3 (%)	.2513	.2429	.1982
	.4338	.4289	.3987
MPCE_Q4 (%)	.2139	.2040	.1326
	.4100	.4030	.3391
Father's Occupation			
Legislators	.0128	.	.0582
	.1123	.	.2342
Professionals	.0182	.	.0227
	.1336	.	.1491
Technicians	.0296	.	.0237
	.1695	.	.1523
Clerks	.0231	.	.0147
	.1502	.	.1202
Services	.0867	.	.0647
	.2813	.	.2459
Agriculture	.3828	.	.3329

	.4861	.	.4713
Craft Workers	.0902	.	.1437
	.2865	.	.3508
Plant Operators	.0410	.	.0548
	.1982	.	.2276
Elementary Occupations	.3157	.	.2845
	.4648	.	.4512
Zone			
North	.2575	.2603	.3382
	.4372	.4388	.4731
South	.2222	.2011	.0992
	.4157	.4008	.2989
East	.2216	.2206	.2556
	.4153	.4147	.4362
West	.1854	.1857	.1657
	.3886	.3889	.3718
Central	.0750	.0875	.1023
	.2634	.2825	.3030
NE	.0384	.0449	.0391
	.1920	.2070	.1940
N	27242	21739	12590

Mixed Families with Unwanted Girls

	(1)	(2)	(3)
	1986-87 (Round 42)	***1995-96 (Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std.Dev.	Std.Dev.	Std.Dev.
No. of children	3.533	3.420	3.380
	.8175	.6988	.7241
No. of girls	2.533	2.420	2.380
	.81755	.6988	.7241
No. of boys	1	1	1
	0	0	0
Father Edu Years	4.472	5.385	6.310
	4.610	4.750	4.749
Mother Edu Years	2.420	2.935	4.263
	3.678	4.126	4.4389
Urban	.2442	.2793	.2714
	.4297	.4488	.4448
SC	.1756	.1725	.1942
	.3806	.3779	.3957
ST	.0782	.0704	.1055
	.2686	.2559	.3073
Others	.7462	.7571	.7003
	.4353	.4290	.4582
HH size	5.891	5.725	5.631
	1.440	1.343	1.135
Female Headed HH	.0647	.0558	.0715
	.2462	.2296	.2577
MPCE (INR)	970.01	1188.06	1435.04
	405.25	501.56	722.67
MPCE_Q1 (%)	.2755	.2280	.3541
	.4469	.4197	.4783
MPCE_Q2 (%)	.2390	.2711	.2157
	.4266	.4447	.4114
MPCE_Q3 (%)	.2128	.2308	.2287
	.4094	.4215	.4201
MPCE_Q4 (%)	.2726	.2700	.2015
	.4455	.4441	.4012
Father's Occupation			
Legislators	.0226	.	.0875
	.1488	.	.2827
Professionals	.0200	.	.0341
	.1400	.	.1815
Technicians	.0411	.	.0235
	.1985	.	.1515
Clerks	.0278	.	.0182
	.1644	.	.1339
Services	.0966	.	.0705
	.2956	.	.2560

Agriculture	.3154	.	.3167
	.4648	.	.4653
Craft Workers	.0965	.	.1497
	.2954	.	.3569
Plant Operators	.0496	.	.0594
	.2172	.	.2364
Elementary Occupations	.3303	.	.2404
	.4705	.	.4274
Zone			
North	.1958	.2339	.2575
	.3970	.4234	.4374
South	.30793	.2892	.1698
	.4618	.4535	.3756
East	.2161	.1593	.1991
	.4117	.3661	.3994
West	.1735	.2107	.2272
	.3788	.4079	.4191
Central	.0770	.0826	.1180
	.2666	.2754	.3227
NE	.0297	.0242	.0284
	.1697	.1539	.1662
N	1700	1913	1906

Source: Authors' calculations based on NSS data

Table 3: Child-level Descriptives

Girls in All-Girls Families

	(1)	(2)	(3)
	1986-87(Round 42)	***1995-96(Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std. Dev.	Std. Dev.	Std. Dev.
Age	10.643	10.444	11.586
	3.499	3.565	3.811
No. of siblings	2.744	2.692	2.587
	.9601	.8921	.9414
Ever_enrolled	.6115	.7120	.9397
	.4875	.4529	.2380
Currently_enrolled	.8376	.9062	.9448
	.3689	.2917	.2284
Years of Education	2.931	3.2457	4.861
	3.006	2.978	2.911
Pvt_school	.2888	.2885	.3339
	.4533	.4532	.4718
Eng_md school	.	.1179	.1958
	.	.3225	.3969
Pvt-Eng_md school	.	.1008	.1500
	.	.3012	.3572
Edu_exp	2145.48	2561.90	5197.97
	2161.30	2693.77	7978.00
Stream of Education			
Arts	.5900	.6539	.4009
	.4986	.4823	.4922
Commerce	.2555	.2100	.3495
	.4422	.4129	.4788
Science	.1545	.1361	.2496
	.3664	.3476	.4346
N	2823	2867	2093

Boys in All-Boys Families

	(1)	(2)	(3)
	1986-87(Round 42)	***1995-96(Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std. Dev.	Std. Dev.	Std. Dev.
Age	12.46	12.552	13.06
	3.885	3.883	3.866
No. of siblings	2.717	2.624	2.462
	.9669	.8990	.7979
Ever_enrolled	.6981	.8193	.9425
	.4591	.3848	.2328
Currently_enrolled	.8529	.8562	.9139
	.3542	.3509	.2806
Years of Education	3.667	4.198	5.464
	3.169	3.024	3.093
Pvt_school	.2633	.2579	.3370
	.4405	.4375	.4728
Eng_md School	.	.0648	.1437
	.	.2462	.3509
Pvt-Eng_md School	.	.0475	.1181
	.	.2126	.3227
Edu_exp	2106.63	2458.90	5043.83
	2030.71	2368.46	7500.83
Stream of Education			
Arts	.4186	.4276	.4388
	.4958	.4978	.4971
Commerce	.2385	.3174	.3331
	.4283	.4683	.4721
Science	.3429	.2550	.2282
	.4771	.4386	.4204
N	11154	8750	4313

Girls in Mixed Families with no Unwanted Girls

	(1)	(2)	(3)
	1986-87(Round 42)	***1995-96(Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std. Dev.	Std. Dev.	Std. Dev.
Age	11.41	11.50	12.49
	3.587	3.655	3.75
No. of siblings	3.578	3.344	3.156
	1.365	1.274	1.243
Ever_enrolled	.5603	.6760	.9121
	.4964	.4680	.2831
Currently_enrolled	.7879	.8348	.9162
	.4088	.3714	.2772
Years of Education	2.835	3.187	5.081
	3.063	3.022	3.130
Pvt_school	.2607	.2448	.2961
	.4390	.4300	.4566
Eng_md School	.	.0650	.1199
	.	.2466	.3249
Pvt-Eng_md School	.	.0456	.0874
	.	.2087	.2825
Edu_exp	2045.17	2334.33	4217.34
	2040.08	2308.59	6394.76
Stream of Education			
Arts	.5768	.5990	.5639
	.4950	.4913	.4962
Commerce	.2467	.2173	.2882
	.4318	.4134	.4532
Science	.1765	.1837	.1479
	.3819	.3882	.3552
N	38489	30066	16584

Boys in Mixed Families with no Unwanted Girls

	(1)	(2)	(3)
	1986-87(Round 42)	***1995-96(Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std. Dev.	Std. Dev.	Std. Dev.
Age	11.75	11.83	12.66
	3.813	3.799	3.810
No. of siblings	3.531	3.253	3.058
	1.383	1.290	1.222
Ever_enrolled	.7249	.7975	.9319
	.4466	.4018	.2520
Currently_enrolled	.8692	.8804	.9173
	.3372	.3246	.2754
Years of Education	3.647	3.803	5.171
	3.036	2.918	3.022
Pvt_school	.2508	.2510	.3520
	.4335	.4336	.4776
Eng_md School	.	.0654	.1485
	.	.2473	.3556
Pvt-Eng_md School	.	.0457	.1158
	.	.2088	.3200
Edu_exp	2040.95	2316.53	5142.18
	2005.84	2317.37	7790.13
Stream of Education			
Arts	.4579	.4089	.4405
	.4990	.4928	.4967
Commerce	.2455	.4119	.4432
	.4310	.4934	.4970
Science	.2967	.1792	.1163
	.4575	.3845	.3208
N	43892	33199	17395

Girls in Mixed Families with Unwanted Girls

	(1)	(2)	(3)
	1986-87(Round 42)	***1995-96(Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std. Dev.	Std. Dev.	Std. Dev.
Age	11.01	11.37	12.23
	3.590	3.663	3.809
No. of siblings	2.796	2.611	2.600
	1.000	.8302	.9173
Ever_enrolled	.6100	.7966	.9666
	.4878	.4026	.1796
currently_enrolled	.8482	.8820	.9533
	.3589	.3226	.2111
Years of Education	3.017	3.970	5.563
	3.099	3.151	3.0458
Pvt_school	.3231	.3246	.3264
	.4678	.4683	.4690
Eng_md School	.	.0821	.1749
	.	.2746	.3800
Pvt-Eng_md School	.	.0666	.1397
	.	.2494	.3467
Edu_exp	2210.95	2687.33	5647.90
	2143.56	2596.65	8743.88
Stream of Education			
Arts	.6590	.6640	.3378
	.4782	.4763	.4739
Commerce	.2347	.1543	.4000
	.4275	.3642	.4908
Science	.1064	.1818	.2622
	.3110	.3889	.4407
N	3469	3861	3452

Boys in Mixed Families with Unwanted Girls

	(1)	(2)	(3)
	1986-87(Round 42)	***1995-96(Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std. Dev.	Std. Dev.	Std. Dev.
Age	9.585	10.04	10.80
	3.286	3.252	3.432
No. of siblings	2.428	2.342	2.280
	.7054	.6682	.5796
Ever_enrolled	.8459	.9467	.9854
	.3613	.2248	.1200
Currently_enrolled	.9449	.9675	.9735
	.2283	.1774	.1606
Years of Education	3.602	3.967	4.756
	2.574	2.296	2.587
Pvt_school	.3111	.3666	.4864
	.4633	.4821	.5000
Eng_md School	.	.1391	.3069
	.	.3462	.4614
Pvt-Eng_md School	.	.1099	.2731
	.	.3129	.4457
Edu_exp	2004.02	2813.47	7869.52
	2184.01	2722.18	10288.15
Stream of Education			
Arts	.1756	0	.1510
	.4394	0	.3617
Commerce	.8244	1	.6169
	.4394	0	.4910
Science	0	0	.2321
	0	0	.4264
N	841	1104	1273

All Boys in the Sample

	(1)	(2)	(3)
	1986-87(Round 42)	***1995-96(Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std. Dev.	Std. Dev.	Std. Dev.
Age	11.86	11.93	12.65
	3.842	3.83	3.828
No. of siblings	3.347	3.102	2.911
	1.346	1.239	1.163
Ever_enrolled	.7213	.8056	.9362
	.4484	.3958	.2443
Currently_enrolled	.8676	.8783	.9197
	.3389	.3270	.2717
Years of Education	3.651	3.888	5.208
	3.057	2.931	3.021
Pvt_school	.2544	.2562	.3564
	.4355	.4365	.4789
Eng_md School	.	.0677	.1560
	.	.2512	.3629
Pvt-Eng_md School	.	.0481	.1246
	.	.2140	.3303
Edu_exp	2053.03	2360.79	5274.78
	2014.23	2343.04	7921.67
Stream of Education			
Arts	.4470	.4109	.4334
	.4978	.4928	.4957
Commerce	.2479	.3915	.4177
	.4323	.4889	.4934
Science	.3051	.1976	.1489
	.4610	.3988	.3561
N	55887	43053	22981

All Girls in the Sample

	(1)	(2)	(3)
	1986-87(Round 42)	***1995-96(Round 52)***	***2014 (Round 71)***
	Mean	Mean	Mean
	Std. Dev.	Std. Dev.	Std. Dev.
Age	11.32	11.40	12.36
	3.588	3.659	3.771
No. of siblings	3.445	3.208	3.016
	1.343	1.237	1.197
Ever_enrolled	.5686	.6920	.9232
	.4953	.4617	.2663
Currently_enrolled	.7980	.8467	.9252
	.4015	.3603	.2632
Years of Education	2.859	3.277	5.135
	3.063	3.042	3.103
Pvt_school	.2696	.2594	.3051
	.4438	.4383	.4605
Eng_md School	.	.0721	.1369
	.	.2586	.3438
Pvt-Eng_md School	.	.0534	.1026
	.	.2249	.3035
Edu_exp	2070.09	2401.10	4561.21
	2061.46	2387.29	7043.24
Stream of Education			
Arts	.5901	.6204	.5036
	.4925	.4861	.5002
Commerce	.2458	.2015	.3161
	.4312	.4018	.4652
Science	.1641	.1781	.1803
	.3709	.3832	.3846
N	44781	36794	22129

Source: Authors' calculations based on NSS data

Table 4: Overall Gender Gaps in Various Educational Outcomes

	(1) **Ever enrol**	(2) **Curr enrol**	(3) **Ed Yrs**	(4) **Pvt**	(5) **Eng**	(6) **PvtEng**
Female	-0.118***	-0.0711***	-0.659***	-0.0169**	-0.0192***	-0.0127***
	(-32.20)	(-17.94)	(-22.80)	(-2.96)	(-4.18)	(-3.34)
1995-96	0.0779***	0.0135**	0.216***	0.00212		
	(13.79)	(2.82)	(6.70)	(0.40)		
2014	0.214***	0.0666***	1.196***	0.120***	0.122***	0.109***
	(37.71)	(13.28)	(34.91)	(18.82)	(28.00)	(28.76)
1995-96 * Female	0.0237***	0.0252***	0.233***	-0.00476		
	(3.76)	(4.25)	(5.29)	(-0.54)		
2014 * Female	0.0855***	0.0733***	0.663***	-0.0362***	-0.00566	-0.0131*
	(9.13)	(9.94)	(14.88)	(-4.08)	(-0.93)	(-2.50)
Father's Edu Years	0.0229***	0.00649***	0.141***	0.00677***	0.00405***	0.00382***
	(46.19)	(14.42)	(55.32)	(12.37)	(10.03)	(10.53)
Mother's Edu Years	0.0212***	0.00947***	0.0589***	0.00907***	0.00673***	0.00568***
	(14.92)	(15.30)	(19.40)	(13.98)	(14.91)	(14.13)
No. of siblings	-0.0176***	-0.00401***	-0.134***	0.000704	0.00203	0.000958
	(-15.73)	(-3.70)	(-16.27)	(0.39)	(1.31)	(0.65)
Urban	0.0281***	-0.00733*	0.105***	0.163***	0.0801***	0.0705***
	(7.88)	(-2.23)	(5.27)	(34.82)	(21.05)	(20.50)
ST	-0.0477***	-0.0109	-0.322***	-0.0282***	0.0362***	-0.00140
	(-9.37)	(-1.77)	(-9.69)	(-3.39)	(6.60)	(-0.30)
SC	0.00172	0.00133	-0.0484*	-0.0512***	-0.0294***	-0.0231***
	(0.47)	(0.38)	(-1.99)	(-10.07)	(-6.80)	(-6.45)
Age	0.00684***	-0.0355***	0.446***	0.0103***	0.000739	-0.00241***
	(15.54)	(-62.72)	(147.76)	(17.24)	(1.51)	(-5.64)
Eldest_sibling	-0.0234***	-0.00361	-0.225***	0.00560	0.0121**	0.0131***
	(-6.16)	(-1.10)	(-9.35)	(1.14)	(2.94)	(3.77)
MPCE Q2	0.0266***	0.0128**	0.203***	0.0396***	0.0317***	0.0224***
	(6.27)	(3.03)	(7.45)	(6.55)	(7.36)	(6.38)
MPCE Q3	0.0634***	0.0255***	0.462***	0.0691***	0.0615***	0.0535***
	(14.90)	(5.90)	(16.77)	(11.44)	(13.88)	(13.23)
MPCE Q4	0.111***	0.0508***	0.864***	0.118***	0.146***	0.123***
	(23.23)	(11.49)	(28.72)	(16.96)	(24.90)	(22.63)

South	0.105***	-0.0386***	0.717***	-0.108***	0.0793***	0.0704***
	(23.43)	(-8.63)	(23.97)	(-16.51)	(13.76)	(13.94)
East	-0.00525	0.00737	-0.0114	-0.211***	-0.0314***	-0.0116*
	(-1.10)	(1.79)	(-0.39)	(-33.39)	(-6.43)	(-2.57)
West	0.0455***	-0.0154***	0.339***	-0.117***	-0.0471***	-0.0243***
	(10.16)	(-3.60)	(12.38)	(-18.73)	(-10.80)	(-6.83)
Central	0.0482***	0.0112	0.200***	-0.228***	-0.0488***	-0.0197***
	(7.69)	(1.88)	(5.18)	(-30.92)	(-10.09)	(-4.30)
North-East	0.0902***	0.0295***	0.608***	-0.282***	0.101***	0.0262***
	(16.25)	(6.03)	(17.20)	(-43.46)	(14.96)	(5.06)
Observations	183126	127695	183412	125229	77986	77696

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Authors' calculations based on NSS data

Table 5: Gender Gaps in Various Educational Outcomes, by Childtype

	(1) **Ever enrol**	(2) **Curr enrol**	(3) **Ed Yrs**	(4) **Pvt**	(5) **Eng**	(6) **PvtEng**
Only Girls	-0.109***	-0.0693***	-0.254**	-0.0386*	0.0134	0.0214
	(-9.22)	(-3.83)	(-2.91)	(-2.28)	(1.01)	(1.62)
Girls_Mixed	-0.0927***	-0.0696***	-0.399***	-0.0307**	-0.0133	-0.0148*
	(-14.10)	(-11.65)	(-6.25)	(-3.15)	(-1.84)	(-2.06)
Boys_Mixed	0.0284***	0.00196	0.331***	-0.0142	0.00613	-0.000459
	(4.45)	(0.37)	(5.07)	(-1.50)	(0.79)	(-0.06)
Girls_Unwanted	-0.107***	-0.0615***	-0.420***	0.00684	-0.0375***	-0.0292***
	(-8.96)	(-5.36)	(-4.35)	(0.38)	(-4.17)	(-3.49)
Boys_Unwanted	0.0622*	0.0328*	0.577***	0.0251	0.00318	-0.00462
	(2.40)	(2.14)	(4.79)	(0.88)	(0.21)	(-0.38)
1995-96	0.0931***	0.000421	0.407***	-0.0113		
	(7.30)	(0.04)	(5.08)	(-0.97)		
2014	0.237***	0.0691***	1.488***	0.0900***	0.117***	0.104***
	(18.25)	(7.13)	(18.28)	(6.34)	(12.91)	(12.21)
1995-96 * Only Girls	0.0300	0.0526*	0.109	0.0453		
	(1.52)	(2.57)	(0.84)	(1.88)		
1995-96 * Girls_Mixed	0.00660	0.0334***	-0.00469	0.00676		
	(0.58)	(3.39)	(-0.05)	(0.47)		
1995-96 * Boys_Mixed	-0.0157	0.0144	-0.236**	0.0167		
	(-1.34)	(1.45)	(-2.70)	(1.19)		
1995-96 * Girls_Unwanted	0.0475**	0.0496**	0.330*	-0.00321		
	(2.67)	(3.12)	(2.56)	(-0.14)		
1995-96 * Boys_Unwanted	0.0124	0.0296	-0.480**	0.0314		
	(0.25)	(0.91)	(-2.85)	(0.81)		
2014 * Only Girls	0.0541*	0.0723**	-0.0176	0.0264	-0.0125	-0.0293
	(2.17)	(3.27)	(-0.14)	(1.03)	(-0.74)	(-1.91)
2014 * Girls_Mixed	0.0503**	0.0614***	0.379***	-0.00224	-0.00540	-0.0102
	(3.03)	(5.25)	(4.29)	(-0.14)	(-0.52)	(-1.05)
2014 * Boys_Mixed	-0.0340*	-0.00748	-0.339***	0.0347*	0.00177	0.00184
	(-2.03)	(-0.63)	(-3.81)	(2.24)	(0.17)	(0.20)
2014 * Girls_Unwanted	0.169***	0.105***	0.536***	-0.0539*	0.0275*	0.0117
	(6.13)	(6.01)	(4.47)	(-2.32)	(2.01)	(0.97)

2014 * Boys_Unwanted	0.0653	0.0112	-1.078***	0.0480	0.0454*	0.0374*
	(1.05)	(0.28)	(-7.29)	(1.41)	(2.44)	(2.40)
Father's Edu Years	0.0229***	0.00643***	0.141***	0.00675***	0.00402***	0.00379***
	(46.13)	(14.30)	(55.30)	(12.32)	(10.00)	(10.51)
Mother's Edu Years	0.0210***	0.00935***	0.0589***	0.00892***	0.00662***	0.00559***
	(14.77)	(15.07)	(19.35)	(13.74)	(14.69)	(13.95)
No. of siblings	-0.0180***	-0.00365**	-0.139***	0.00138	0.00232	0.00151
	(-15.59)	(-3.28)	(-16.60)	(0.76)	(1.49)	(1.02)
Urban	0.0277***	-0.00749*	0.103***	0.163***	0.0799***	0.0703***
	(7.79)	(-2.29)	(5.18)	(34.91)	(21.09)	(20.60)
ST	-0.0481***	-0.0111	-0.324***	-0.0281***	0.0360***	-0.00150
	(-9.46)	(-1.80)	(-9.77)	(-3.38)	(6.52)	(-0.32)
SC	0.00193	0.00167	-0.0470	-0.0507***	-0.0292***	-0.0229***
	(0.53)	(0.48)	(-1.95)	(-9.96)	(-6.80)	(-6.40)
Age	0.00711***	-0.0353***	0.447***	0.0105***	0.000934	-0.00224***
	(16.18)	(-62.44)	(148.84)	(17.50)	(1.90)	(-5.21)
Eldest_sibling	-0.0251***	-0.00481	-0.246***	0.00652	0.0131**	0.0139***
	(-6.55)	(-1.45)	(-10.12)	(1.31)	(3.15)	(3.94)
MPCE Q2	0.0262***	0.0126**	0.202***	0.0399***	0.0318***	0.0225***
	(6.18)	(2.98)	(7.43)	(6.60)	(7.39)	(6.42)
MPCE Q3	0.0631***	0.0254***	0.463***	0.0693***	0.0610***	0.0531***
	(14.84)	(5.90)	(16.86)	(11.48)	(13.82)	(13.20)
MPCE Q4	0.111***	0.0506***	0.866***	0.118***	0.146***	0.123***
	(23.35)	(11.48)	(28.83)	(16.95)	(24.83)	(22.60)
South	0.103***	-0.0404***	0.712***	-0.109***	0.0774***	0.0685***
	(23.10)	(-8.95)	(23.87)	(-16.69)	(13.43)	(13.61)
East	-0.00531	0.00737	-0.0101	-0.211***	-0.0316***	-0.0118**
	(-1.12)	(1.80)	(-0.35)	(-33.40)	(-6.47)	(-2.60)
West	0.0447***	-0.0162***	0.335***	-0.117***	-0.0475***	-0.0247***
	(10.03)	(-3.78)	(12.28)	(-18.73)	(-10.87)	(-6.91)
Central	0.0471***	0.0108	0.199***	-0.228***	-0.0494***	-0.0201***
	(7.53)	(1.83)	(5.20)	(-30.95)	(-10.27)	(-4.39)
North-East	0.0894***	0.0290***	0.606***	-0.282***	0.101***	0.0259***
	(16.15)	(5.94)	(17.22)	(-43.44)	(14.85)	(5.02)
Observations	183126	127695	183412	125229	77986	77696
t statistics in parentheses						

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Authors' calculations based on NSS data

Table 6: Educational Expenditure: Heckman two-step estimation

	(1) **Girls**	(2) **Child Type**
Edu_exp		
Female	256.7***	
	(9.61)	
1995-96	281.5***	401.8***
	(9.72)	(6.02)
2014	2259.7***	2093.5***
	(27.71)	(11.71)
Female * 1995-96	-68.66	
	(-1.54)	
Female * 2014	-765.0***	
	(-6.78)	
Only Girls		547.6***
		(5.87)
Girls_Mixed		340.4***
		(7.77)
Boys_Mixed		134.7**
		(3.19)
Girls_Unwanted		514.5***
		(6.61)
Boys_Unwnated		354.6**
		(2.68)
Only Girls * 1995-96		-26.70
		(-0.16)
Only Girls * 2014		-392.5
		(-1.38)
Girls_Mixed * 1995-96		-223.2**
		(-2.94)
Girls_Mixed * 2014		-869.1***
		(-4.38)
Boys_Mixed * 1995-96		-176.0*
		(-2.37)
Boys_Mixed * 2014		-15.82

		(-0.08)
Girls_Unwanted * 1995-96		-154.0
		(-1.26)
Girls_Unwanted * 2014		211.9
		(0.72)
Boys_Unwnated * 1995-96		381.1
		(1.92)
Boys_Unwnated * 2014		2657.1***
		(5.75)
Age	371.7***	380.0***
	(43.06)	(44.01)
Constant	-1875.7***	-2076.4***
	(-20.01)	(-20.45)
Enrolment status for children between 6-16 years		
Female	-0.494***	
	(-19.69)	
1995-96	0.0702*	-0.0138
	(2.56)	(-0.24)
2014	0.508***	0.522***
	(13.29)	(7.30)
Female * 1995-96	0.182***	
	(4.67)	
Female * 2014	0.492***	
	(9.61)	
Only Girls		-0.508***
		(-4.88)
Girls_Mixed		-0.479***
		(-11.33)
Boys_Mixed		0.00896
		(0.21)
Girls_Unwanted		-0.499***
		(-7.24)
Boys_Unwnated		0.185
		(1.15)
Only Girls * 1995-96		0.360**

		(2.62)
Only Girls * 2014		0.510**
		(3.29)
Girls_Mixed * 1995-96		0.245***
		(3.77)
Girls_Mixed * 2014		0.422***
		(5.19)
Boys_Mixed * 1995-96		0.108
		(1.65)
Boys_Mixed * 2014		-0.0263
		(-0.32)
Girls_Unwanted * 1995-96		0.382***
		(3.71)
Girls_Unwanted * 2014		0.850***
		(6.42)
Boys_Unwnated * 1995-96		0.219
		(1.05)
Boys_Unwnated * 2014		0.393
		(1.27)
Father's Edu Years	0.0486***	0.0481***
	(15.62)	(15.50)
Mother's Edu Years	0.0793***	0.0784***
	(16.84)	(16.58)
No. of siblings	-0.0323***	-0.0300***
	(-4.54)	(-4.10)
Urban	0.00784	0.00651
	(0.34)	(0.29)
ST	-0.0989*	-0.102**
	(-2.54)	(-2.64)
SC	-0.00742	-0.00613
	(-0.31)	(-0.26)
Age	-0.219***	-0.218***
	(-47.22)	(-47.11)
Eldest_sibling	-0.00780	-0.0155
	(-0.36)	(-0.70)

MPCE Q2	0.136***	0.136***
	(5.23)	(5.24)
MPCE Q3	0.257***	0.258***
	(9.35)	(9.38)
MPCE Q4	0.557***	0.557***
	(15.99)	(16.01)
South	-0.262***	-0.275***
	(-9.56)	(-9.97)
East	0.0267	0.0287
	(0.91)	(0.97)
West	-0.111***	-0.118***
	(-3.89)	(-4.10)
Central	0.0176	0.0154
	(0.41)	(0.36)
North-East	0.192***	0.187***
	(4.92)	(4.80)
Constant	3.363***	3.343***
	(47.93)	(43.59)
athrho		
Constant	-0.375***	-0.381***
	(-18.72)	(-18.08)
lnsigma		
Constant	8.336***	8.332***
	(472.98)	(473.05)
Rho	-.3583	-.3634
Sigma	4173	4153.24
Lamba	-1496	-1509.3
Observations	116517	116517

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Author's calculations based on NSS data

Table 7: Birth Spacing, in years

		Birth Spacing in Years (Mean)		
		(1)	(2)	(3)
		1986-87	***1995-96***	***2014***
Between 1st and 2nd				
All Girls		3.46	3.03	3.12
All Boys		4.64	4.37	3.80
Mixed	Boy	4.37	4.09	3.66
	Girl	3.04	2.87	2.74
Unwanted	Girl	3.21	2.90	2.87
Between 2nd and 3rd				
All Girls		3.22	3.03	3.11
All Boys		4.12	3.70	3.34
Mixed	Boy	3.46	3.30	2.99
Girl	2.90	2.80	2.79	
Unwanted	Girl	3.04	3.08	3.05
Between 3rd and 4th				
All Girls		1.10	1.09	0.87
All Boys		1.37	1.06	0.76
Mixed	Boy	2.21	1.79	1.50
	Girl	1.67	1.44	1.20
Unwanted	Girl	2.77	2.74	2.88
Between 4th and 5th				
All Girls		0.83	0.78	0.98
All Boys		1.04	0.98	0.86
Mixed	Boy	1.46	1.42	1.14
	Girl	1.31	1.14	1.08
Unwanted	Girl	2.60	2.60	2.41
Between 5th and 6th				
All Girls		0.80	0.67	0.73
All Boys		0.81	0.72	0.58
Mixed	Boy	1.18	1.03	1.00
	Girl	1.06	0.84	1.01
Unwanted	Girl	2.98	3.09	3.49
(Note: In "unwanted" families, the boy is the last child.)				

Source: Authors' calculations based on NSS data

Figures

Figure 1: Gender Gaps in the Probability of Ever Enrolling

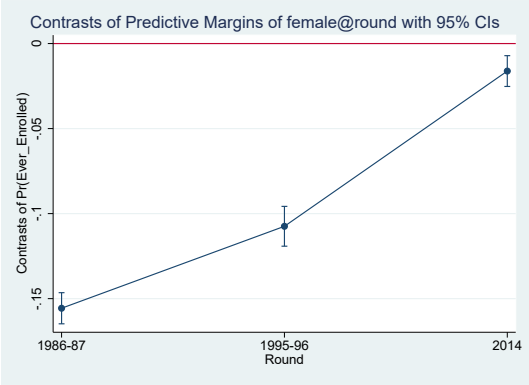


Figure 2: Gender Gaps in the Probability of Current Enrolment

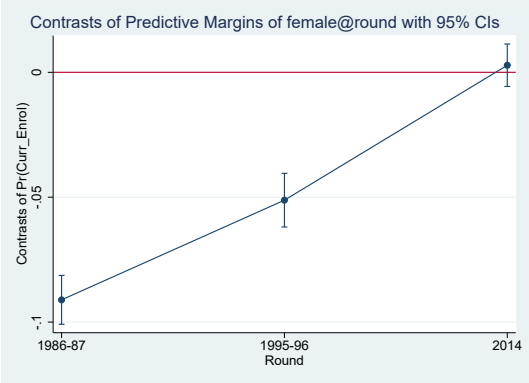


Figure 3: Predicted Probability of Ever Enrolling, by Childtype

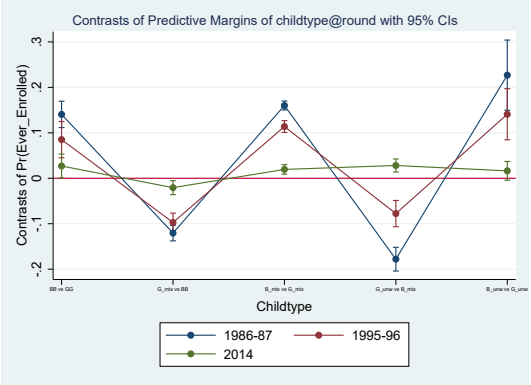


Figure 4: Predicted Probability of Current Enrolment, by Childtype

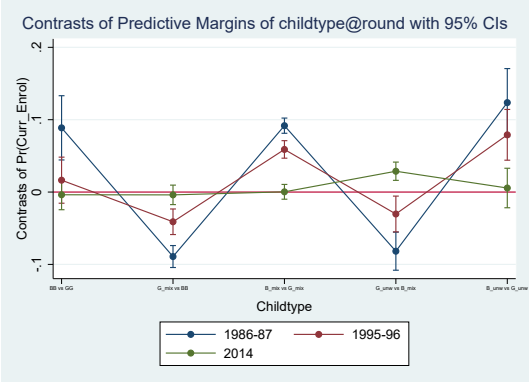


Figure 5: Gender Gaps in Years of Education

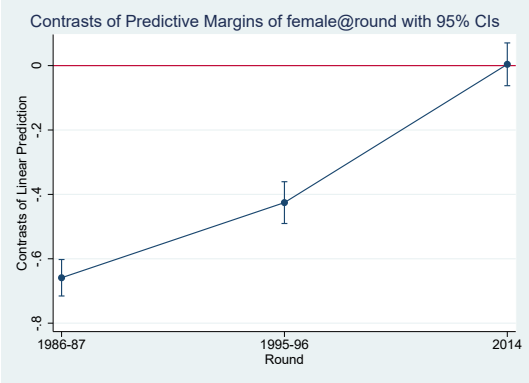


Figure 6: Gender Gaps in Years of Education, by Childtype

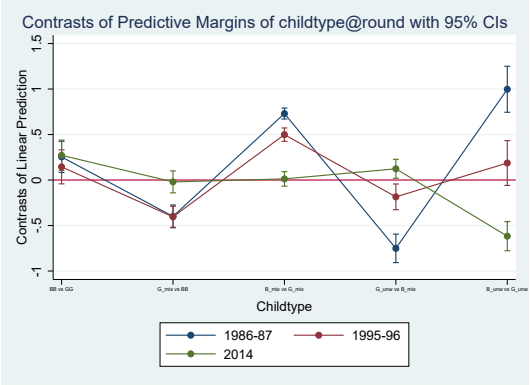


Figure 7: Gender Gaps in Private Schooling

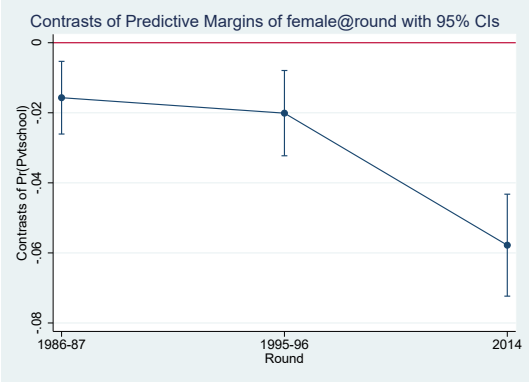


Figure 8: Gender Gaps in English Medium Schooling

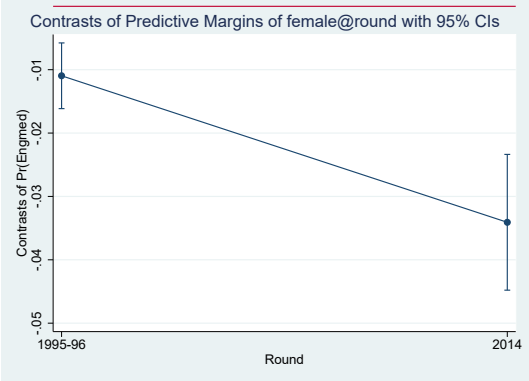


Figure 9: Gender Gaps in Private Schooling, by Childtype

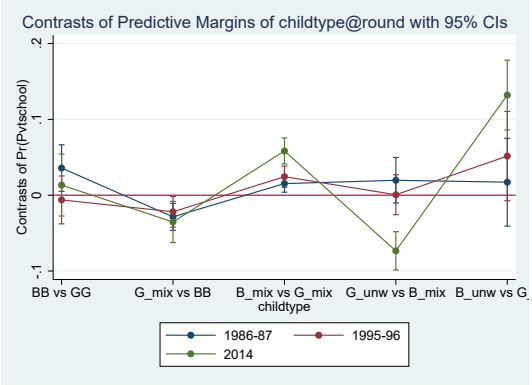


Figure 10: Gender Gaps in English Medium Schooling, by Childtype

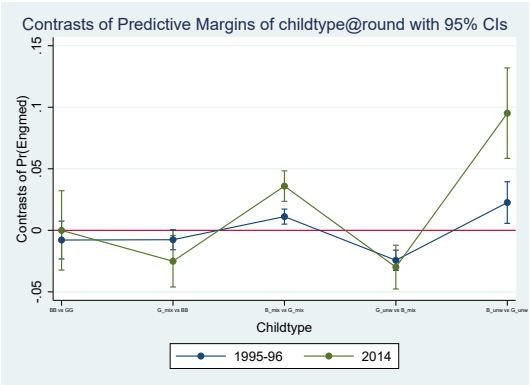


Figure 11: Gender Gaps in Edu Expenditure s.t. Selection

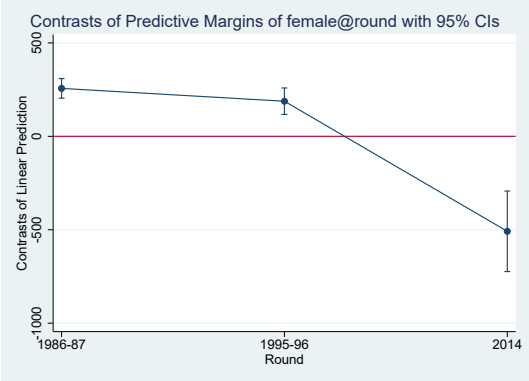


Figure 12: Gender Gaps in Edu Expenditure s.t. Selection, by Childtype

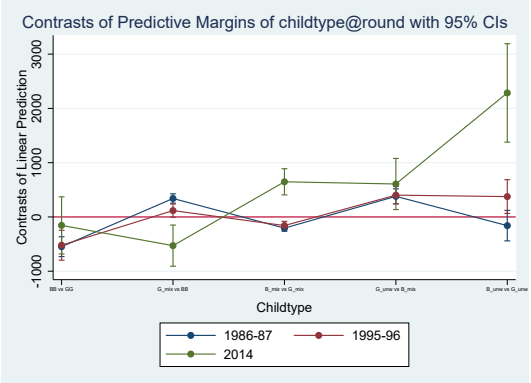


Figure 13: Gender Gaps in Stream Choice

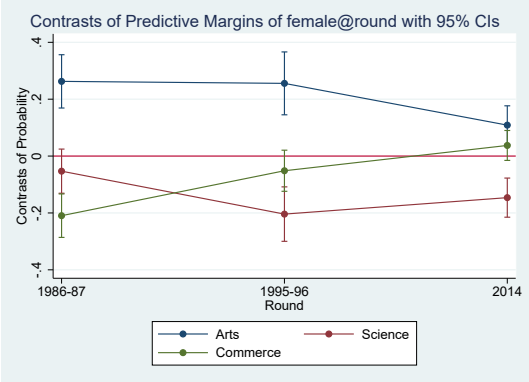
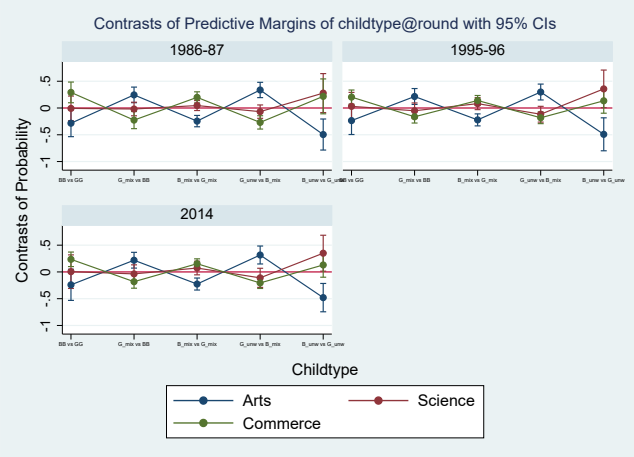


Figure 14: Gender Gaps in Stream Choice, by Childtype



References

[1] Anderson, S., and Ray, D. (2010), "Missing women: age and disease", *The Review of Economic Studies*, 77(4), 1262-1300.

[2] Attané, I., and Guilmoto, C. Z. (2007), "Watering the neighbour's garden: The growing demographic female deficit in Asia"

[3] Avinash, T. (2017), "A study on Karnataka state government initiatives and schemes for promoting primary education"

[4] Azam, M., Chin, A., and Prakash, N. (2013), "The returns to English-language skills in India", *Economic Development and Cultural Change*, 61(2), 335-367

[5] Basu, A. M., and Desai, S. (2016), "Hopes, dreams and anxieties: India's one-child families", *Asian Population Studies*, 12(1), 4-27

[6] Bongaarts, J. (2013), "The implementation of preferences for male offspring", *Population and Development Review*, 39(2), 185-208

[7] Bongaarts, J., and Guilmoto, C. Z. (2015), "How many more missing women? Excess female mortality and prenatal sex selection, 1970–2050", *Population and Development Review*, 41(2), 241-269

[8] Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009), "Women's underrepresentation in science: sociocultural and biological considerations", *Psychological bulletin*, 135(2), 218

[9] Chung, W., and Gupta, M. D. (2007), "The decline of son preference in South Korea: The roles of development and public policy", *Population and Development Review*, 33(4), 757-783

[10] Dahl, G. B., and Moretti, E. (2008), "The demand for sons", *The Review of Economic Studies*, 75(4), 1085-1120

[11] Das Gupta, M., and Mari Bhat, P. N. (1997), "Fertility decline and increased manifestation of sex bias in India", *Population Studies*, 51(3), 307-315

[12] Das Gupta, M., Zhenghua, J., Bohua, L., Zhenming, X., Chung, W., and Hwa-Ok, B. (2003), "Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea", *The Journal of Development Studies*, 40(2), 153-187

[13] Deaton, A. (1989), "Looking for Boy-Girl Discrimination in Household Expenditure Data", *The World Bank Economic Review*, 3(1), 1-15

[14] Desai, Sonalde; Dubey, Amaresh; Vanneman, Reeve; and Banerji, Rukmini (2009), "Private schooling in India: A new educational landscape", in: Bery, S.; Bossworth, B.; Panagariya, A., editors. *India Policy Forum*. Vol. 5. New Delhi: Sage Publications

[15] Deshpande, A. and Nordman, C.J. (2018), "Gender Discrimination in Schooling: The Role of Institutions, Infrastructure and Gender Norms", *mimeo*

[16] Deshpande, A. and Ramachandran, R. (2019), "Traditional hierarchies and affirmative action in a globalizing economy: Evidence from India", *World Development*, Vol 118, 63-78

[17] Fors, Congdon H., and Lindskog, A. (2017), "Within-Family Inequalities in Human Capital Accumulation in India: Birth Order and Gender Effects"

[18] Government of India (2018), *Economic Survey 2017-18*, Volume I, Ministry of Finance, Department of Economic Affairs, Economic Division, India, 2018

[19] Jayachandran, S. (2015), "The roots of gender inequality in developing countries", *Economics*, 7(1), 63-88.

[20] Jayachandran, S. (2017), "Fertility decline and missing women", *American Economic Journal: Applied Economics*, 9(1), 118-39.

[21] John, M. E., Kaur, R., Palriwala, R., Raju, S., and Sagar, A. (2008), "Planning families, planning gender: The adverse child sex ratio in selected districts of Madhya Pradesh, Rajasthan, Himachal Pradesh, Haryana and Punjab", *ActionAid*, New Delhi, India

[22] John, Mary (2018), *The Political and Social Economy of Sex Selection: Exploring Family-Development Linkages*, UNFPA and UNWomen

[23] Kabeer, N., Huq, L., and Mahmud, S. (2014), "Diverging stories of 'missing women' in South Asia: Is son preference weakening in Bangladesh?", *Feminist Economics*, 20(4), 138-163

[24] Kaur, R., Bhalla, S. S., Agarwal, M. K., and Ramakrishnan, P. (2016), "Sex ratio at birth. the role of gender, class and education", UNFPA technical report, 2016: 43

[25] Kaur, R., and Vasudev, C. (2019), "Son Preference and Daughter Aversion in Two Villages of Jammu", *Economic and Political Weekly*, 54(13), 13

[26] Kingdon, G.G. (2005), "Where has all the bias gone? Detecting gender bias in the intrahousehold allocation of educational expenditure", *Economic Development and Cultural Change*, 53(2), 409-451

[27] Kingdon, G.G. (2007), "The progress of school education in India", *Oxford Review of Economic Policy*, 23(2), 168-195

[28] Kishor, S., and Gupta, K. (2009), "Gender equality and women's empowerment in India", *National Family Health Survey (NFHS-3)*, India 2005-06

[29] Lancaster, Geoffrey, Pushkar Maitra and Ranjan Ray (2008), "Household Expenditure Patterns and Gender Bias: Evidence from Selected Indian States", *Oxford Development Studies*, 36:2, 133-157

[30] Lee, J. (2008), "Sibling size and investment in children?s education: An Asian instrument", *Journal of Population Economics*, 21(4), 855-875

[31] Michael, M., King, L., Guo, L., McKee, M., Richardson, E., and Stuckler, D. (2013), "The mystery of missing female children in the Caucasus: an analysis of sex ratios by birth order", *International perspectives on sexual and reproductive health*, 97-102

[32] Milazzo, A. (2014), *Son preference, fertility and family structure: Evidence from reproductive behavior among Nigerian women*, The World Bank

[33] Mills, M., and Begall, K. (2010), "Preferences for the sex-composition of children in Europe: A multilevel examination of its effect on progression to a third child", *Population Studies*, 64(1), 77-95

[34] Muralidharan, K., and Prakash, N. (2017), "Cycling to school: Increasing secondary school enrollment for girls in India", *American Economic Journal: Applied Economics*, 9(3), 321-50

[35] Munshi, K. and Rosenzweig, M. (2006), "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy", *The American Economic Review*, 96(4), 1225-1252

[36] Rai, R. K. (2017), "Decline in contraceptive use in India: A call for action", *Sexual and reproductive healthcare*, 14, 13-16

[37] Sahoo, S. (2017), "Intra-household gender disparity in school choice: Evidence from private schooling in India", *The Journal of Development Studies*, 53(10), 1714-1730.

[38] Sankar, D. (2008), "What is the Progress in Elementary Education Participation in India During the Last Two Decades?: An Analysis Using NSS Education Rounds", *World Bank*

[39] Sekher, T. V. (2012), "Ladlis and Lakshmis: financial incentive schemes for the girl child", *Economic and Political Weekly*, 58-65

[40] Sen, A. (1990), "More than 100 million women are missing", *New York*, 61-66.

[41] Subramanian, S., and Deaton, A. (1991), "Gender effects in Indian consumption patterns", *Sarvekshana*, 14(4), 1-12.

Appendix

Explicit Son Preference

We see some important differences in family characteristics across the four family types. Could it be the case that we are misinterpreting differences in characteristics of these families as differences in their son preference or parental motivation attitudes?

The actual number and gender composition of children is to some extent determined by parental preferences, but is equally a matter of luck and random events. Thus, in principle, families with only girls could also have a deep-seated son preference, but were simply unlucky and ended up with girls instead of boys. Based on their preferences for an optimal family size, they stopped after a certain number of children, even when they had only girls. Thus, observing the ex-post outcome of births, i.e. only girls, does not indicate anything about whether these families have no/low son preference, i.e. happy with having only girls, or whether they have high son preference, but were simply unlucky.

NSS data do not contain any information on the underlying preferences of parents about the gender composition of their children. Another nationally representative, large scale survey, The National Family Health Survey (NFHS)[17] has an explicit set of questions eliciting son preference. There are three rounds of NFHS that overlap with the time period of our NSS data: 1992-93 (NFHS 1), 1998-99 (NFHS 2) and 2005-06 (NFHS 3). NFHS data asks few questions on son preference to ever-married women in the age group of 15-49 years in 1992-93 (Round 1) and 1998-99 (Round 2), and to all women and men in the age group of 15-49 years in 2005-06 (Round 3). Exactly similar to what we did with NSS data, we identified all the biological children in a family in the age group of 6-19 years, and then divided the families into only girls families, only boys families and mixed families with and without unwanted girls.

NFHS asks two broad questions which allow us to capture son preference among families. First is the ideal number of boys and girls that a woman would like to have.[18] The tabulation in Tables A1 and A2 shows that regardless of family type, i.e. the actual gender composition of children, mothers would like to have similar number of boys and girls. Though on average, they prefer to have more boys over girls, this does not vary across family types. For men, this question was asked only in 2005-06. Again there is no difference across family types in the number of boys and girls they prefer in a family. Second, NFHS asks "Would you prefer your next child to be a boy or a girl or doesn't it matter?" Table A3 shows that a significantly higher proportion of mothers in GG families prefer their next child to be a boy, rather than a girl, which could reflect an underlying son preference. Based on NFHS data, it appears that GG families do not necessarily represent lower son preference, and that they might just have been "unlucky". Thus, parents of only girls are not fundamentally different in unobservable attitudes compared to parents in mixed families. There seems to be near universal instrumental son preference, but ex-post, girls in families with unwanted girls suffer the consequences of that son preference.

[17] NFHS is the Indian module of the Demographic and Health Survey (DHS) data, and is a nationally representative household survey data set on topics related to health, nutrition, domestic violence, fertility decision, son preference, and many other indicators.

[18] The exact set of questions are: "If you could choose exactly the number of children to have in your whole life, how many would that be? How many of these children would you like to be boys, how many would you like to be girls and for how many would the sex not matter?"

Table A1: Ideal Number of boys, by family type

	(1)	(2)	(3)	(4)	(5)
	Only Girls	***Only Boys***	***Mixed***	***Unwanted***	***Total***
Ideal Number of Boys by Mothers					
1986-87					
Mean	1.42	1.57	1.55	1.46	1.54
Std. Dev.	0.78	0.93	0.89	0.82	0.89
1995-96					
Mean	1.43	1.71	1.71	1.34	1.66
Std. Dev.	0.79	1.01	0.94	0.76	0.93
2014					
Mean	1.25	1.46	1.54	1.21	1.48
Std. Dev.	0.69	0.90	0.84	0.71	0.84
Ideal Number of Boys by Fathers					
1986-87					
			Not Asked		
1995-96					
			Not Asked		
2014					
Mean	1.14	1.39	1.42	1.11	1.37
Std. Dev.	0.70	1.00	0.96	0.78	0.94

Source: Authors' calculations based on NFHS data

Table A2: Ideal Number of girls, by family type

	(1) *Only Girls*	(2) *Only Boys*	(3) *Mixed*	(4) *Unwanted*	(5) *Total*
	Ideal Number of Girls by Mothers				
1986-87					
Mean	1.06	1.06	1.05	1.06	1.05
Std. Dev.	0.68	0.63	0.66	0.73	0.66
1995-96					
Mean	1.13	0.98	1.17	1.07	1.14
Std. Dev.	0.70	0.60	0.71	0.65	0.69
2014					
Mean	1.04	0.92	1.10	1.00	1.07
Std. Dev.	0.62	0.54	0.65	0.60	0.63
	Ideal Number of Girls by Fathers				
1986-87					
	Not Asked				
1995-96					
	Not Asked				
2014					
Mean	0.97	0.85	0.98	0.89	0.96
Std. Dev.	0.66	0.57	0.71	0.66	0.69

Source: Authors' calculations based on NFHS data

Table A3: Son Preference through NFHS, by family type

	Son Preference for next child by Mothers (%)				
	(1)	(2)	(3)	(4)	(5)
	Only Girls	***Only Boys***	***Mixed***	***Unwanted***	***Total***
1986-87					
Boy	44.66	48.15	50.83	43.56	49.66
Girl	6.52	7.95	11.58	12.33	10.87
Doesn't matter	23.96	28.6	25.82	26.71	26.16
Depends upon God	24.87	15.31	11.78	17.41	13.31
Total	100	100	100	100	100
1995-96					
Boy	97.48	12.72	62.69	89.63	70.76
Girl	0	74.87	6.5	0.81	9.66
Doesn't matter	0.84	4.72	14.42	3.96	8.92
Depends upon God	1.68	7.69	16.39	5.6	10.66
Total	100	100	100	100	100
2014	**Not Asked**				
	Son Preference for next child by Fathers (%)				
1986-87					
Boy	57.73	56.14	51.98	44.35	52.35
Girl	4.37	9.2	10.5	8.06	9.87
Doesn't matter	22.56	22.61	24.08	32.06	24.31
Depends upon God	15.34	12.05	13.44	15.53	13.47
Total	100	100	100	100	100
1995-96	**Not Asked**				
2014	**Not Asked**				

Source: Authors' calculations based on NFHS data