ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Maitra, Pushkar; Pal, Sarmistha; Sharma, Anurag

Working Paper

Absence of Altruism? Female Disadvantage in Private School Enrolment in India

IZA Discussion Papers, No. 9808

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Maitra, Pushkar; Pal, Sarmistha; Sharma, Anurag (2016) : Absence of Altruism? Female Disadvantage in Private School Enrolment in India, IZA Discussion Papers, No. 9808, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/141567

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 9808

Absence of Altruism? Female Disadvantage in Private School Enrolment in India

Pushkar Maitra
Sarmistha Pal
Anurag Sharma

March 2016

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Absence of Altruism? Female Disadvantage in Private School Enrolment in India

Pushkar Maitra
Monash University

Sarmistha Pal
University of Surrey
and IZA

Anurag Sharma
Monash University

Discussion Paper No. 9808
March 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 9808
March 2016

ABSTRACT

Absence of Altruism? Female Disadvantage in Private School Enrolment in India*

Using two nationally representative datasets from household surveys conducted in India in 2005 and 2012, the present paper examines the causal effect of gender in private school choice. We argue that the gender of the child is potentially endogenous in India because parents continue to have children until they have a son. To redress this potential endogeneity, we exploit the variation in private school choice among 7-18 year olds born to the same parents within the same household in an attempt to minimize both child-invariant and child-varying household-level omitted variable bias. We then explore the nature of female (dis)advantage across different types of households, communities and years with a view to assess the role of parental preferences in this respect and also its change. Significant female disadvantage exists in both survey years, though the size of this disadvantage varies across sub-samples and years. Female disadvantage is significantly higher among younger (relative to eldest) girls and also in northern and northwestern (relative to western) regions, but it is lower among girls from poor (relative to rich) households, Christian (relative to Hindu high caste) households, and those with more educated mothers. We infer that the observed within-household variations in female disadvantage across sub-samples reflect variations in non-altruistic parental preferences linked to deeply held cultural norms (for example, sons acting as old-age security and the exogamy of girls), access to schools and other public goods, thus posing considerable challenges to securing 'education for all.'

JEL Classification: C21, I25, O10

Keywords: private school enrolment, female disadvantage, parental altruism, household fixed-effects model, India

Corresponding author:

Sarmistha Pal
Surrey Business School
University of Surrey
Guildford GU2 7XH
United Kingdom
E-mail: s.pal@surrey.ac.uk

* Sarmistha Pal gratefully acknowledges the funding provided by the Leverhulme Research Fellowship, which initiated this project. We are very grateful to the editor and the three anonymous referees for their constructive comments on earlier versions of the paper. We have benefitted from discussions with Marco Alfano, Alpaslan, Akay, Mark Bray, Tarun Jain, Jyotsna Jalan, Dilip Mookherjee, Janneke Pieters, Andrew Oswald, Debraj Ray, Zahra Siddique, Francis Teal, and the participants in the IZA staff seminar series, Oxford Conference on Education and Development, Growth and Development Workshop at the Indian Statistical Institute Kolkata, and IZA-World Bank Annual Conference on Employment and Development for their helpful feedback at various stages. The usual disclaimer applies.

1. Introduction

Women make up one-half of the world's population. Without their engagement, empowerment, and contributions, we cannot hope to achieve rapid economic growth or effectively tackle global challenges such as climate change, technology adoption, food security, and conflict. However, two-thirds of world's non-literate adults are women, a damning statistic that highlights the pervasive denial of the human right to education experienced by women and girls across the globe. Education can empower women, justifying the need to secure 'education for all,' an essential component of the Millennium Development Goals. It is therefore important to understand what limits girls' schooling. While there is a sizeable literature on gender discrimination in girls' schooling (see, for example, CEDAW (2012)), we know little if anything about girls' access to private schooling, despite its recent rapid growth around the world. Naturally, our knowledge of gender discrimination in schooling will remain incomplete without a good understanding of the nature and extent of gender discrimination in private schooling. In an attempt to bridge this gap in the literature, we examine the role of gender in private school choice.

Gender is likely to have an ambiguous effect on private schooling. On one hand, private school growth may lower the extent of the gender gap in schooling: first, private schools may be more suited to the needs of girls if, for example, they ensure local access to schools, as well as to more female teachers; second, parents who choose private schools for their children are likely to be more motivated and altruistic and hence less likely to discriminate between boys and girls. On the other hand, private schools require fees and thus may increase the extent of discrimination against girls in a country in which parents rely on sons for old-age security and dowries, exogamy, and female disadvantage in labour markets are pervasive, reducing the value of investments in girls. We use 2005 and 2012 India Human Development Survey (IHDS) data to examine the nature, size, and variations of female disadvantage across households, communities and also time, if any, in private school choice in India, with a view to drawing out the implications in terms of parental preferences and possible policy interventions.

Despite impressive growth in income and literacy over the last two decades, India continues to lag behind the other BRIC countries in indices of enrolment and attendance, especially beyond the primary level of schooling (Kingdon (2007). According to the World Economic Forum's Global Gender Gap Report 2009, India ranked 114th out of 134 countries measured. Recent data and studies (see, for example, Drèze and Sen (2013)) highlight the worsening human development situation and increasing gender inequity since the 1990s, [1] a period that has also interestingly witnessed the considerable growth of private schools around the country.[2] Despite the absence of school fees, the dismal state of government-run schools has induced many households, even some poorer ones, to take advantage of the newly emerging private unaided schools in India to meet their educational needs. This has also been facilitated by the modest private school fees in India (Tooley and Dixon (2003)).

We use individual-level unit record data from 7–18 year olds from two rounds of nationally representative India Human Development Survey (IHDS) conducted in 2005 and 2012. We argue that in societies characterised by a strong preference for sons, the gender of the child is endogenous. This is because in such societies, parents may continue to have children until that have the preferred gender composition (see Kishor (1993)). The latter has been worsened by the availability (since the early 1980s) of mechanisms (e.g., scanning technology) that enable sex-selective abortions (see Jha, et al. (2011)). In such an environment, the same unobserved parental characteristics that may affect the gender of a child could also systematically affect the educational opportunities of boys and girls differently, thus causing significant endogeneity bias in school choice estimates.

[1] There are worrying statistics showing that the Gender Inequality Index (GII) for India has worsened between 2008 and 2011. India now ranks 129 out of 146 countries on the GII, better only than Afghanistan within South Asia (Times of India).

[2] While about 16% of the villages surveyed in 1996 as a part of the PROBE data used by Drèze and Kingdon (2001) had access to private schools, the corresponding figure rose to about 28% in 2003 (see Muralidharan and Kremer (2006)). Our computations using the Indian Human Development Survey data show that 24 and 31% of children aged 7–18 attended private school in 2005 and 2012, respectively.

While some use the gender of the first child on the grounds that it is random, thereby restricting the analysis to the first-born (see, for example, Rosenzweig and Wolpin (2000)),[3] we use a household fixed-effects approach that exploits the variation in the school choice for children born to same parents. This approach not only addresses the endogeneity of the gender of the child but also the issue of potential bias arising from the omitted household-level child-invariant (for example, parental motivation or lack of it) and child-varying (for example, individual ability and/or behavioural traits) factors that may also influence their school choice. This allows us to estimate the causal effect of gender on private school enrolment.

Daughters may receive less human capital investment than sons (regardless of market returns) if parents inherently place a low value on females.[4] In the context of India, there is evidence of significant discrimination against girls for a range of different outcome variables: fertility (Kishor (1993)), malnutrition (Pal (1999)), school enrolment and educational attainment (Pal (2004)), and breastfeeding (Jayachandran and Kuziemko (2011)). This observed gender inequity could be the result of intra-household allocation rules (Browning and Chiappori (1998)), cultural norms (sons as old-age security or exogamy (Dyson and Moore (1983)), and/or state interventions influencing the supply of relevant public goods, including schools (Murthi, et al. (1995)). Similar gender-based differences are observed in other developing countries as well (Glick (2008)). Economists have generally attributed it to gender differences in the labor market in terms of models of intra-household allocation in which lower female agency arises from either lower productivity or lower returns from female labor (Kingdon (1996); Sundaram and Vanneman (2008)). Further human capital investment in girls has been shown to increase when work opportunities requiring more education arise (Heath and Mobarak (2011) and also when women have better information about and access to existing jobs (Jensen (2012)).

[3] Note, however, that the randomness of the gender of the first child can also be questioned given the widespread use of scanning technology since the early 1980s.
[4] See Das Gupta (1987); Behrman (1988); Davies and Zhang (1995); Kingdon (2002).

In comparison, the literature on gender-based discrimination in private schooling is rather under-developed in both its scope and methodology, especially in the Indian context.[5] Using a sample of low-income households in the state of Uttar Pradesh in India, Srivastava (2006) finds no evidence of gender inequity in private school enrolment. When using 13-village survey data from the same state of Uttar Pradesh in India, Härmä (2009) observes that a son is preferred for private education, especially for secondary schools, which is then attributed to household budget constraints. Using two rounds of Young Lives longitudinal data from the Indian state of Andhra Pradesh, Woodhead, et al. (2013) report a widening gender gap in low-fee private school enrolment over the 2001–2009 period. Unlike the present paper, all of these cited papers have used data from small and localized surveys and do not control for the endogeneity of gender and other household-level omitted-variable bias.[6]

Our study builds on various strands of the existing literature in an attempt to identify the causal effect of gender on private schooling, with implications regarding its nature, size, causes, and evolution in an all-India context. We estimate private school enrolment as a function of a child's gender (if female, as captured by the GIRL dummy), birth order, and dummies representing age categories, with controls for child-invariant and child-varying unobserved household-level factors within a household fixed-effects model. In our framework, the estimated coefficient of the GIRL dummy yields the marginal effect of being a girl (relative to a boy) child which we take to be a measure of female disadvantage in private school choice. We assess its variation not only across different types of households (by caste, religion, expenditure, parental education) and communities (rural/urban, northern, north-western, western regions of the country), but also over time 2005-2012. Our results indicate that there is a significant bias against

[5] The existing literature has primarily focused on the relative efficiency of public and private schools, as measured by the effect of school type on various student test scores. Both positive and negative effects on the part of private schools are found (e.g., see Chudgar and Quin (2012)). Building on the case for the greater efficiency of private schools, there have been attempts by policy makers to explore the scope of the private sector in delivering basic education in India (see Tooley and Dixon (2003)).

[6] While Woodhead, et al. (2013) have access to panel data, they do not exploit the panel dimension of the data and focus instead on pooled analysis.

girls in private school enrolment and that this persisted for both the 2005 and 2012 rounds of IHDS using matched household panel. The average female disadvantage was about 4 percentage points in 2005, and this rose to about 6 percentage points in 2012. Indeed, estimates from matched 2005 and 2012 panel households do not indicate any drop in female disadvantage over the period. We also find substantial heterogeneity across various sub-samples characterised by individual, household, and community characteristics: female disadvantage in school choice is higher among younger (relative to older) girls but lower for Christian (relative to high caste Hindu) girls and also those from poor (relative to rich) households in which the mothers have at least secondary schooling. There is also interesting regional variation in female disadvantage. One way of interpreting this is that household preferences are generally non-altruistic in an environment in which intra-household allocation rules, job market considerations, and other deeply-rooted socio-cultural norms influence parental schooling decisions as boys and girls grow into adolescents.

We thus contribute to the existing literature in a number of ways. Noting that the net effect of gender could be ambiguous, we exploit the variation in school choice among children born to same parents to identify the causal effect of gender using a household fixed-effects model. Second, many recent surveys ask direct questions about parental (especially women's) preferences, e.g., whether they value girl's schooling equally to boys. Answers to these questions are largely subjective and prone to measurement errors. Instead, we compare the nature and size of female disadvantage across various sub-samples defined by individual/household/community characteristics. In doing so, we incorporate a number of socio-cultural and economic explanations for gender inequity with a view to deriving the nature of parental preferences and distinguishing altruism from misogyny and rationality. This is an important exercise because ultimately, the effectiveness of policies aimed at improving the condition of women in developing countries is crucially dependent on the nature of parental preferences. Policymakers must take this into account.

2. Background, Data and Selected Descriptive Statistics

There are three broad types of recognized schools in India, namely government schools, private aided schools, and private unaided schools. Government and private aided schools are typically government-recognized; i.e., they have the government's stamp of approval. They are similar in many respects because private aided schools are almost entirely financed by the government and have little control over staffing (hiring/firing decisions) and fees, despite being nominally privately managed.

<INSERT TABLE 1 HERE>

Table 1 compares main characteristics of private unaided, government, and private aided schools in 1992 and 2002.[7] The biggest difference between private and government schools is in terms of infrastructure and pupil-teacher ratios. As of 2002, 71% of the private schools had a toilet, as compared to 41% of government schools and 86% of private aided schools; 91% of private school has drinking water facilities, as compared to 78% of government schools and 91% of private aided schools. The pupil-teacher ratio in government schools is often twice that in private schools. All of this suggests that private schools offer better facilities than government schools in India. Furthermore, compared to government schools, a greater proportion of teachers in private schools are women and are likely to be of a higher caste. This is possibly because as of 2002, the private educational sector was not constrained by caste-based affirmative action (reservation) policies operating in the Indian public sector. A comparison of government schools with private aided schools suggests that the latter tend to have similar pupil-teacher ratios to those of government schools. Although private aided schools had poorer physical infrastructure in 1992, they caught

[7] An anonymous referee suggested that we update this information using more recent data. The alternatives are the All India School Education Survey (AISES) data or the District Information System for Education (DISE) data. Unfortunately, the 2009 AISES data are not yet publicly available. The publicly available DISE data only have information on the number of pupils per teacher at the primary school level by school management type. As of 2010-11, the pupil teacher ratios are 29, 31, and 31 for private unaided, government, and private aided schools, respectively. No information on teacher characteristics and school infrastructure by school management type is publicly available. Additionally the AISES and the DISE data are not directly comparable: the DISE data are a sample, while the AISES data are a census of all recognized schools. Further, when making comparisons over time, using the DISE data is problematic because over the years, the sampling framework has changed.

up with the government schools in 2002. In our analysis, we do not distinguish between government and private aided schools; instead, we combine them under the broad umbrella of government schools.[8] Private unaided schools (whether recognized or not), in contrast, enjoy more autonomy as compared to private aided schools and are typically self-funded out of fee income. Thus, private unaided schools are the truly private schools in India.

The present paper uses data from two waves (2005 and 2012) of the Indian Human Development Survey (IHDS). This is a nationally representative, multi-topic survey of 41,554 households in 1,503 villages and 971 urban neighbourhoods across India collected by the National Council of Applied Economic Research and the University of Maryland. Around 83% of the households were re-interviewed in 2011-2012 for the second wave of data, and the response rates were more than 90% for both waves. The survey collected information about health, education, employment, economic status, marriage, fertility, gender relations, and social capital.

<INSERT FIGURE 1 HERE>

We focus on children aged 7-18 years at the time of the survey. Figure 1 presents the average enrolment rates by age and gender for the two waves of data.[9] Several observations are worth noting. First, the enrolment rates are almost 100% for children aged 7-11; they start falling by age 11 and decrease significantly for both boys and girls by age 18. Second, while there is no gender difference in enrolment rates in either wave for children aged 7-10, beyond the age of 10 (i.e., once children have completed primary school), gender differences become the norm. For children aged 11-18, boys are systematically more likely to be enrolled in school. These trends are consistent across the two waves of data, though in absolute terms, the enrolment levels for both boys and girls are higher in 2012. Third, the drop in enrolment rates for the 15-18 year olds

[8] As of 2005, 66.11% of children were enrolled in a government school, 27.53% in a private school, and 6.36% in a private aided school or religious school. The corresponding percentages in 2012 were 58.78, 34.58, and 6.64, respectively. We exclude religious schools from our analysis.

[9] In Figure A1, we present the trends in ever having attended school by gender. The pattern, in all cases, follows an inverted u-shape. In 2005, 90% of all children in the 7-18 age group had ever attended school; the corresponding figure had increased to 96% in 2012, indicating an increase in ever having attended school.

is sharp, dropping to 35% and 20% for boys and girls, respectively, in 2005 and to 50% and 40%, respectively, in 2012. This drop may reflect boys leaving schools in search of employment and girls leaving school because they have attained marriageable age. Overall, 69% of boys in the 15-18 age group who are not enrolled in school are likely to be engaged in market work, as compared to 0.5% of boys enrolled in school and 17% of girls not enrolled in school. Girls in the 15-18 age group who are not enrolled in school are more likely to be married: 19.4% of girls aged 15-18 and not enrolled in school are likely to be married, as compared to 0.8% of enrolled girls, 0.6% of enrolled boys, and 3.4% of non-enrolled boys. Table A1 in the Appendix summarizes these arguments regarding children aged 15-18 at the time of the 2012 survey.

<INSERT FIGURE 2 HERE>

Figure 2 presents the average private school enrolment by age and gender, conditional on any enrolment, for both waves. Table 2 presents the averages for several sub-samples. Overall, Figure 2 and Table 2 both reveal interesting patterns relating to private school enrolment for boys and girls. The 2005-2012 period saw a systematic increase in private school enrolment rate for both boys and girls: private school enrolment rates increased from 27% to 36% for boys and from 23% to 29% for girls over the survey years. However, the gender difference in private school enrolment increased from 4.1 percentage points (or 17.6% of the mean private school enrolment rate for girls in 2005) to 6.7 percentage points (or 23.3% of the mean private school enrolment rate for girls in 2012). There is a change in the pattern of gender bias across age groups. For children aged 7-14, the average difference in private school enrolment between boys and girls was about 5 percentage points in 2005 and increased to 7.7 percentage points in 2012. For children aged 15-18, while there was no gender difference in private school enrolment in 2005, the gap increased to 3.7 percentage points in 2012. While the gender gap is statistically significant for the 15-18 year olds in 2012, it is considerably lower than that for children aged 7-14.

<INSERT TABLE 2 HERE>

The pro-male bias in private school enrolment persists, irrespective of sector of residence, household expenditure quartile, religion, caste (the exception being Christians – see Columns 5

and 10), whether one's parents are more educated or not (though the difference is not statistically significant in 2005 for children born to mothers with more than secondary schooling), and whether one resides in the northwest, north, east or south of the country.[10]

<INSERT TABLE 3 HERE>

Table 3 presents the averages for the variables used in our analysis. There are remarkable consistencies across the two waves. In both surveys, 48% of children are females, and 31% of children reside in urban areas; these are most likely to be Hindu SC/ST/OBC. Mothers are, on average, considerably less educated than fathers. There are increases in the rates of ever having attended school and private school enrolment. Per capita household expenditure (in nominal terms) increased almost two-and-a-half-fold to Rs. 1700 across the two waves. The female-to-male wage ratio at the community level (primary sampling unit) dropped slightly, from 73% to 71%, over this period. The ratio of the average distance from the nearest government school to the nearest private school increased slightly, from 1.4 to 1.9.

3. Empirical Strategy, Identification, and Hypotheses

3.1. Econometric Issues and Empirical Strategy

The primary focus of this paper is on school choice and, in particular, private school enrolment of 7-18-year-old Indian children. We aim to identify the causal effect of gender on private school enrolment after addressing the potential sources of estimation bias. First, estimation bias may arise from the potential endogeneity of the child's gender. This is particularly pertinent in societies with a strong preference for sons. Second, the non-enrolment of children continues to be prevalent in the Indian context, which may give rise to selectivity bias. Selectivity bias is primarily a reflection of omitted household (e.g., parental support and motivation) and/or individual (e.g.,

[10] We divide the country into six different regions. Northwest (consisting of Jammu and Kashmir, Himachal Pradesh, Punjab, and Haryana), north (consisting of Uttarkhand, Delhi, Uttar Pradesh, Rajasthan, Chhattisgarh, and Madhya Pradesh), east (consisting of Bihar, Jharkhand, and West Bengal), northeast (consisting of Assam, Meghalaya, Arunachal Pradesh, Manipur, Mizoram, Nagaland, and Tripura), West (consisting of Goa, Gujarat, and Maharashtra), and South (consisting of Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu).

ability, health issues, or behavioural problems) characteristics (Heckman (1979)). The problem gets worse when these unobservables are correlated with the residual error term, which is particularly difficult to tackle in cross-sectional or pooled data. To address these potential estimation biases, we use a household fixed-effects model in which the estimation sample is restricted to households with multiple children aged 7-18 years. This allows us to exploit the intra-household variation in private school enrolment among children born to the same parents and thus identify the *causal* effect of gender on private school enrolment. Accordingly, this analysis focuses on the characteristics of individual children (gender, age categories, and birth order) that vary within households after eliminating household-level omitted factors (both child-invariant and child-varying).

Define S_{ij}^{*} as the propensity of the ith child from the jth household to enrol in a private unaided (henceforth private) school at the time of the survey. This propensity is determined by the following equation:

$$S_{ij}^{*} = \beta' X_{ij} + \eta_j + v_{ij} \tag{1}$$

Here, S_{ij}^{*} (the propensity of the child to attend private school) is not observed; what we observe instead is binary variable S_{ij}. This equals 1 the ith child from the jth household is enrolled in a private school at the time of the survey and 0 if he or she attends a government or private aided school. In Equation (1), η_j accounts for the household-level child-invariant unobserved characteristics, and v_{ij} is the IID error term.

Addressing selectivity bias is particularly problematic in single cross-section or pooled data analysis because many household- and individual-level unobserved factors remain unaccounted for. An obvious solution, in this case, is to make use of the Heckman-type selection models that attempt to account for the unobserved ability bias in various ways. Note, however, that there is no general selection model available for panel data, which is important in the execution of a household fixed-effects approach such as ours (with the exception of Wooldridge (1995) and Semykina and Wooldridge (2013), which only work for balanced panels thus lose a

large number of observations).[11] We consider several alternatives. We start with the OLS variant of Equation (1) to determine the school choice of the ith child in the jth household, using household fixed effects within a panel framework.[12] This approach accounts for unobserved child-invariant household-level characteristics (e.g., parental motivation, support, and other family circumstances), thus minimizing the omitted-variable bias on this account.

These estimates may still be biased because of the presence of unobserved within-household child-varying factors. These could include unobserved child ability or behavioral traits, which may influence that child's school choice. This, in turn, could play an important role in influencing parental decisions regarding both enrolment and school choice. In order to address this omitted-variable bias,[13] we estimate an extended version of Equation (1) that also includes unobserved traits varying across different children (in ascending birth order) within the same household. The regression specification is given by the following:

$$S_{ij}^{*} = \beta' X_{ij} + \eta_j + \left(\theta_i \times \eta_j\right) + v_{ij} \tag{2}$$

Specification (2) is an improvement over Specification (1) because it accounts for the unobserved child-invariant and child-varying factors within the household.[14] It is often argued that sample

[11] Using the 2005 IHDS data, Maitra, et al. (2013) used a variant of Wooldridge (1995) and Semykina and Wooldridge (2013) to estimate private school enrolment among children born to household heads. The methodology required us to estimate the model only for a balanced panel for households with 2, 3, or 4 children, thus losing a substantial number of the sample observations. The results were, however, similar to those presented in the current paper.

[12] There are, of course, a number of reasons a binary logit or probit, not OLS, should be the preferred estimation method. See http://www.mostlyharmlesseconometrics.com/2012/07/probit-better-than-lpm. However the problem is that the marginal effects are not well defined when one uses a non-linear estimation technique such as a logit or probit. Here, marginal effects are crucial because we are not only interested in the direction of the bias but also in its magnitude. Given this constraint, we present the OLS estimates with household fixed effects.

[13] Note that while we can observe the parent-perceived ability of children in the 2005 survey, we do not have this information in 2012. Also, parent-perceived ability is a subjective measure and may suffer from measurement error. Hence, we do not include this variable, even for the 2005 survey results, and instead rely on within-household individual trends.

[14] Note that children are ordered by birth, so our household fixed-effects approach is similar to the notion of a linear time trend in the usual panel data, where child-specific index i for the jth household is replaced with the time dimension t. This approach allows us to account for birth-order-specific unobservables within the household. Taken together, we not only control for unobserved household-level child-invariant (common to all children) fixed effects but also unobserved household-level child-varying trends across children born to same parents at different times. We believe that the inclusion of both these fixed effects will allow us to minimize any estimation bias arising from unobserved heterogeneity.

selection bias can be viewed as a specification error arising from omitted variables. Thus, the inclusion of unobserved household-level child-invariant and child-varying factors in our fixed-effects model is likely to further minimise the potential for any selection bias.

Given that an issue of selection arises for private school choice among those who have been enrolled in any school, it is also interesting to explore how our baseline household fixed effects estimates would compare to a model with selection correction. Because the standard Heckman-type selection model is not suitable for a household fixed-effects set-up, we proceed as follows. School choice is conditional upon school enrolment, so school choice (S_{ij}) is observed only if the ith child from the jth household is enrolled in school at the time of the survey $(E_{ij} = 1)$. In estimating Equation (1), we must therefore account for a selection problem in which the selection equation is defined in terms of the decision to send the child to school (irrespective of school type) and is given by the following:

$$E_{ij}^{*} = \gamma' W_{ij} + \zeta_{ij} \tag{3}$$

Here, E_{ij}^{*} (the propensity to attend school) is not observable. We only observe E_{ij}, where $E_{ij} = 1$ if $E_{ij}^{*} > 0$; $E_{ij} = 0$ otherwise. Equation (3) allows us to estimate the inverse Mill's ratio (λ_{ij}) for each sample child, treating the data as pooled (i.e., without accounting for household fixed effects), which is then included as an additional argument in the school choice equation (Equation (1)). Although not ideal, our objective here is to test the robustness of our preferred estimates of Equation 1 and Equation 2. At the second stage, we estimate a school choice equation within a household fixed effect framework as follows:

$$S_{ij}^{*} = \beta' X_{ij} + \eta_j + \lambda_{ij} + v_{ij} \tag{4}$$

Given that the second stage not only includes the individual-level selection correction terms λ_{ij} generated from the first stage but also the child-invariant household fixed-effects terms η_j, it becomes somewhat comparable to the household FE with individual-level trends in Equation (2). Accordingly, this variant of the selection model is likely to have lower omitted-variable bias

relative to the standard Heckman-type selection models using cross-sectional/pooled data. Given the fixed-effects formulation at the second stage, we cannot include any household-level factors, so the within-household variation in child outcomes is solely explained by the individual characteristics that vary within households.

3.2. Hypotheses

In this subsection, we develop hypotheses pertaining to the possible role of various individual characteristics in private school choice.

Gender: Despite some evidence to the contrary (for example, Behrman (1988) argues that Indian parents are generally averse to inequality among children), the notion of son preference among resource-constrained parents in India is now well-accepted. In all our regressions, we include a dummy for a female child (GIRL) to account for gender-based differences, if any, in private school enrolment among siblings born to same parents. In the absence of any prior evidence regarding the role of gender in school choice among children born to same parents, our results remain data-driven.

Birth order: Leaving aside the gender dimension, parents may choose to discriminate between sons and daughters in other ways as well. Garg and Morduch (1998) argue that children are better off in terms of measured health indicators if they have all sisters and no brothers. This is because parents tend to allocate fewer resources to girls. Das Gupta (1987) finds that in rural Punjab, girls with older sisters suffer the most. We include a binary variable indicating whether the index child is the eldest in the family with a view to testing whether the eldest child enjoys any premium in private schooling. Pande and Malhotra (2006) argue that younger girls are worse off than elder girls in India. The underlying idea is that the eldest child may get preferential treatment from his or her parents in terms of their schooling decisions because he or she is more likely to start earning earlier than the other children, thus supplementing the family earnings. By estimating the household fixed-effects model for the sample of non-eldest children, we examine whether younger girls are more discriminated against (relative to the eldest child).

Age: Our data description in Section 2 (see Figure 2 and Table 2) highlights that the gender gap in private school enrolment varies with the age of the child: it is high for 7-9 and 10-14 year olds (4.4 percentage points and 5.4 percentage points in 2005, going up to 7.8 percentage points and 7.6 percentage points, respectively, for these age groups in 2012); note that the gap disappears for the 15-18 year olds in 2005, while it persists but is considerably lower than that of the full sample in 2012. Accordingly, we examine whether children in certain age categories receive preferential treatment from parents regarding their schooling choices. This may highlight the differential costs and benefits of private schooling for male and female children, as they grow older. The primary benefit of schooling comes from earnings, while there are a range of costs, including those of school fees, books, uniform, transport (including security concerns), and also opportunity costs (foregone returns from participating in the labour market because of alternative activities). Both girls and boys are more likely to drop out of school beyond the age of 15. The reason for boys may be that they start supplementing the family income, whereas for girls, personal safety and security issues start becoming important at that age, and this adds to the costs of travelling to and from schools.[15] Additionally girls are likely to get married at this age. While the central and state governments in India have succeeded in expanding access to primary schools such that over 95% of villages have a primary school, access to secondary schools is still rather limited. This may encourage some parents to choose local private schools for girls if they are sent to school at all, especially when they are adolescents.[16] In addition to local access to schools, another factor that may encourage adolescent girls' private school enrolment is the

[15] There is now evidence from a number of countries that supports the argument that reduced distance to schooling has significant effects on the school enrolment of girls. See, for example, Andrabi, et al. (2008), Burde and Linden (2013), and Qureshi (2013).

[16] The issue of local access is important. Long journeys to and from school put girls at additional risk, regardless of age, race, class, caste, or location, because of the potential threat of rape, sexual harassment, intimidation and teasing (see, for example, UN (2000)). The fact that someone may need to accompany a girl on her walk to and from school creates a larger burden associated with girls' schooling. In other words, access to local private schools with less travelling time may increase the likelihood of girls' private schooling. A greater distance to private (relative to government schools) is likely to increase the cost of attending private school for girls (relative to boys). This, in turn, may explain the female disadvantage in private school enrolment. In other words, access to local private schools for girls may close the gender gap in private school enrolment.

regular attendance of teachers in private, but not state, schools (Chaudhury, et al. (2006)), which not only improves learning but also enhances the security of girls within school premises.

Our primary estimating equation is given by the following:

$$S_{ij}^{*} = \beta_0 + \beta_1 GIRL_i + \beta_2 Eldest_i + \sum_k \alpha_k I(Age = k) + \eta_j + v_{ij}; k = 8, \dots, 18 \qquad (1')$$

Here, η_j captures the child-invariant household fixed effects.

To account for within-household child-specific trends, we estimate the following extended version of Equation (1'):

$$S_{ij}^{*} = \beta_0 + \beta_1 GIRL_i + \beta_2 Eldest_i + \sum_k \alpha_k I(Age = k) + \eta_j + (\theta_i \times \eta_j) + v_{ij} \qquad (2')$$

The estimated coefficient of the GIRL dummy (β_1) in Equations (1') and (2') can be interpreted as a measure of female (dis)advantage. If $\beta_1 < 0$, we argue that females are disadvantaged. On the other hand, if $\beta_1 > 0$, the bias is against boys. We estimate equations (1') and (2') for the full sample and also separately for a number of sub-samples. We compare the size and significance of β_1 across various sub-samples characterised by individual/household/community characteristics. This incorporates common arguments explaining gender inequity to identify the nature of parental preferences[17] and distinguishes misogyny ($\beta_1 < 0$) from altruism/rationality (insignificant or positive β_1).[18] In the absence of any prior information about this, we let our data speak for themselves.[19]

[17] Parental attitudes towards son preference are sometimes assessed by asking direct questions about parental attitudes towards the sex composition of the ideal family. The 2005 survey provides some direct information as to whether mothers value boys' and girls' schooling equivalently (no information was available for the father). While this information is likely to be subjective and was not available for the 2012 survey, we find a close correspondence between mothers who value boys' and girls' schooling similarly and mothers' levels of education. Only 3% of sample mothers with a son preference had at least 10 years of schooling, while the proportion is as high as 13% for mothers who equally value boys' and girls' schooling. We expect a direct correspondence between parental characteristics and the extent of son preference in schooling.

[18] Altruism refers to a parental aversion to inequality among boys and girls. Rationality, on the other hand, may arise from the fact that there are clear incentives to invest in girls' private schooling, e.g., supplementing family earnings, protecting girls from sexual harassment, and/or ensuring a higher-quality son-in-law.

[19] We estimate a third specification, which is a selection-corrected version of Equation (1'), as in Equation (4). Here, we include a selection correction term λ_{ij}, estimated from a first-stage enrolment equation.

We consider the individual children within a household and apply fixed effects, and in such a case, each household can be considered as a cluster. Fixed effects are an example of modelling "error components," where every observation within the household is correlated with every other observation because children are all genetically linked within a household. We must partial out the fixed effects to derive the homoscedastic idiosyncratic error. This is done by using clustered robust standard errors with household fixed effects (Nichols and Schaffer (2007)).

To examine whether there are statistically significant differences in gender bias across the independent sub-samples, we use a test proposed by Clogg, et al. (1995) and Paternoster, et al. (1998) and used by Fujiwara (2015). Consider regressions on two independent sub-samples, and let β_1 and β_2 be the estimated coefficients from the two sub-samples for the same explanatory variable. Then, the Z test for the difference between the two regression coefficients is given by the following:

$$Z = \frac{\beta_1 - \beta_2}{\sqrt{SE(\beta_1)^2 + SE(\beta_2)^2}} \tag{5}$$

where $SE(\beta_1)$ and $SE(\beta_2)$ are the standard errors of β_1 *and* β_2, respectively, which we compare across complementary sub-samples.[20]

4. **Results**

In this section, we present the results corresponding to all households that were surveyed in each round (2005 and 2012) and also those surveyed in both years (the matched household-level panel).[21] We start with the 2012 results in Section 4.1, which are then compared with the corresponding 2005 results in Section 4.2. Section 4.3 examines the robustness of our results (discussed in Sections 4.1 and 4.2) as compared to alternative samples. Section 4.4 compares the

[20] If there are only two possible groups, the comparison is done for one group versus the other. If there are more than two possible groups, we choose one group as the reference group and compare the coefficient of each sub-sample coefficient with the chosen reference group.

[21] Note, however, that the matched panel is at the household level. Within this set of households, we consider children aged 7-18 in each survey year. The 7-year gap between the two survey rounds implies a small common window at the individual level for this age group.

robustness of the household fixed-effects estimates presented in Sections 4.1–4.3 with the corresponding household fixed-effects estimates with selection correction. Finally, Section 4.5 presents the 2005 and 2012 matched household panel results. All standard errors are clustered at the household level to minimise the autocorrelation of errors within households. Note that our primary estimating sample consists of children aged 7-18 born to the household head and residing in the household at the time of the survey. As a robustness check, we consider an extended sample consisting of all children resident in the household at the time of the survey, irrespective of whether they were born to the household head or not. As we show in Section 4.3, the results for this extended sample are very similar to those for the sample of children born to the household head.

4.1. 2012 Estimates

We start with the household fixed-effects estimates of Equation (1') for 2005 and 2012 rounds of the IHDS. These results are presented in Table 4: Columns 1-4 show the estimates for 2012, while columns 5-8 show those for 2005.

<INSERT TABLE 4 HERE>

In addition to the full sample, we present results corresponding to a number of sub-samples because we attempt to account for the nature of parental preferences in private school choice via observable individual and household characteristics. In this respect, we incorporate the known determinants of gender inequity, including economic and socio-cultural arguments, as well state policy interventions, all of which may influence access to schools and are likely to vary across states. While household fixed-effects estimates control for unobserved household-level, child-invariant characteristics (as in Equation (1')), we use the variations in the coefficient estimates across these sub-samples to account for the nature of parental preferences in private school choice. Thus, in the regression results presented in Table 4 (and all subsequent tables), each row presents the results from a different sub-sample. All regressions not only control for the gender of the child (as reflected in the inclusion of the GIRL dummy) but also control for the age of the child (by including age dummies) and birth order (by including a dummy for the eldest

child). Ceteris paribus, gender bias is captured via the coefficient estimate of the GIRL dummy. Given that these are estimates from fixed effects OLS regressions, the estimated coefficient of the GIRL dummy also gives us the marginal effect, i.e., the size of the female disadvantage in private school enrolment.

There is a large and statistically significant female disadvantage in both the overall sample of 7-18 year olds (row 1) and also separately for 7-14 (row 2) and 15-18 (row 3) year olds. The marginal effects associated with the GIRL dummy show that for the full sample, the GIRL child is, on average, about 6.3 percentage points less likely to attend private school. The corresponding likelihood is similar for the 7-14 year olds but slightly lower (about 4.4 percentage points lower) for the 15-18 year olds. The lower average inequity, as compared to the 7-14 age group, among 15-18-year-old girls may reflect parental concerns about protecting these adolescent girls, which increases their likelihood of being in a private school. Note, however, that the difference in the GIRL coefficient in the 15-18 year age group (relative to 7-14 year olds) is not statistically significant (see column 4).

Splitting the sample by rural (Row 4) or urban (Row 5) residence, we find that the coefficient estimate of the GIRL dummy is slightly higher for the urban sample (at 6.6 percentage points) than for the rural sample (at 6.1 percentage points). Note, however, that the two coefficient estimates are not significantly different, as reflected by the test of equality.

Row 6 shows the female disadvantage among non-eldest children within households. The coefficient estimate of the GIRL dummy is not only negative and statistically significant, but its magnitude is also greater than that of the sample of eldest children. Furthermore, the coefficient estimate of the GIRL dummy is not statistically significant for the sub-sample of eldest children (see Row 7). This difference in the GIRL coefficient between the eldest and the non-eldest subsamples is also statistically significant, as reflected by the Z test of the equality of the GIRL coefficient in Column 4. Taken together, these results suggest that eldest girls are significantly better off than younger girls, a result that is consistent with the findings of Pande and Malhotra (2006).

***Role of household/parental characteristics*:** Rows 8-20 presents the results for sub-samples defined by specific household characteristics: household expenditure quartiles Q1-Q4 (Rows 8-11), household religion and caste (Hindu High caste, Hindu Scheduled caste/Scheduled tribe/Other backward caste, Muslims, and Christians: Rows 12-16), and the educational attainment of the father and mother (Rows 17-20).

Becker and Lewis (1965) argue that investment in children increases at higher levels of income, which has empirical support (see, for example, Pal (1999) and Filmer (2005)). In contrast, the need to supplement family earnings by reaping the benefits from investment in girls' schooling could be lower among the rich (relative to the poor). Thus, the total effect could be ambiguous. The results presented in Rows 8-11 show that the extent of female disadvantage in private school enrolment is lower among poor (those in Q1 and Q2 of the expenditure distribution) households. The results of the Z-test for equality across the expenditure quartiles shows that the extent of gender bias is significantly lower for households in the first quartile as compared to richer fourth quartile households. This perhaps highlights the relatively higher need of the poorer households to supplement family earnings via boys and girls rather than practising discrimination.

We next consider the role of religion in the gender gap in private school enrolment and obtain estimates of female disadvantage for Hindu (upper and lower castes in Rows 12 and 13, respectively), Muslim (Row 14), and Christian (Row 15) households. Within Hindu households, there are no systematic and significant differences in female disadvantage between higher- and lower-caste households because the relevant Z-statistic remains insignificant. Female disadvantage in private school enrolment is, however, somewhat lower for Muslim households (relative to high-caste Hindu households), and the GIRL dummy is not statistically significant in Christian households. Unlike Hindu households, endogamy tends to be more prevalent among Muslims (Korson (1971)) and Christians (Mitterauer (1991)), so Muslim and Christian parents tend to

expect to have stronger lifelong relationships with their daughters (even after marriage), thus explaining the lower female disadvantage among Muslim and Christian households.[22]

Finally, we assessed the effect of parental educational attainment on private school enrolment among girls. Parental preferences may not always be aligned. For example, mothers may have more empathy for their daughters, while fathers may have more empathy for their sons. We classify parents by whether or not they had at least 10 years of schooling (see Rows 17-20). While the GIRL dummy is statistically significant in each case, the female disadvantage is lower when the mother is better educated relative to when she is not (4.3 percentage points as opposed to 6.6). The magnitude of the extent of the gender bias is lower when mothers (fathers) are more educated (relative to when they are less educated) though the Z-statistic is not statistically significant. Although female disadvantage persists, the favourable role of a mother's schooling is consistent with the existing literature on intra-household allocation that argues that a mother's education, a major driver of women's status within households, is crucial for the human capital accumulation of the next generation, especially girls (see, for example, Schultz (2001)).

Role of community characteristics: Next, we examine the role of selected community characteristics, namely school accessibility and the potential returns of private schooling, which are identified as possible reasons for the gender gap in schooling.

Accessibility: Long journeys to and from school place girls at additional risk, and often, an adult must accompany a girl to school. This creates a larger burden associated with girls' schooling. This is one area in which the state can effectively intervene by building new schools and/or providing pubic transport to/from schools, which may improve access to schools, thus reducing inherent female disadvantage in this respect. In our analysis, relative distance is captured by a binary variable that takes a value of 1 when the ratio of the average distance to the public and private school within the district is in the fourth quartile and a value of 0 otherwise. In other

[22] Note, however, that the Z-scores on the sub-sample equality tests, as summarized in Column 4, confirm that the null hypothesis that the effects are equal across Muslims and Christians (relative to Hindu high castes) cannot be rejected.

words, the variable takes a value of 1 when the relevant private school is closer than the closest public school.

Returns to Schooling: Within intra-household allocation models, lower female agency arises from either lower productivity or lower returns from female labour because these models take account of the contributions of each member of the household. In this context, we examine whether and how the observed gender bias is affected by the potential returns of schooling. Duraisamy (2002) finds evidence for the presence and persistence of gender differences the in returns of schooling in India. Kingdon (1998) and Kingdon and Theopold (2006) argue that lower female school participation in India is significantly linked to lower returns for female schooling. If returns for schooling are higher for boys, ceteris paribus, parents may choose fee-paying private schools only for boys with a view to boosting family earnings. In other words, one can expect a reversal of gender discrimination against girls in private school enrolment only when relative expected returns for girls' schooling are higher. Empirically, we generate the female-to-male hourly wage rate for those who have ever attended a school in the primary sampling unit (which we consider to be the immediate community of the child). We prefer this to the individual-level wage earnings because the latter are likely to suffer from potential endogeneity when determining private school enrolment. In particular, we generate a binary variable FM_wage_1, which takes a value of 1 if the ratio of female-to-male market hourly wage in the community exceeds 1, i.e., when labour market returns for women are higher than that for men. The variable takes a value of 0 otherwise. All descriptive statistics are summarised in Table 3.

The results presented in Rows 21-24 show that the bias against the GIRL child persists, even when the cost of attending private school (in terms of travel time and accessibility) is low. While female disadvantage continues to be statistically significant, we find that higher returns and closer private schools reduce the extent of female disadvantage (compare the marginal effects in Rows 21 and 23 to those in Row 1). Note that the extent of gender bias is greater when labour

market returns for women are lower. However, the test of equality between each pair of complementary samples rejects any significant difference in the gender gap between girls living close/far from the school and also those facing high/low female (relative to male) wages.

***Regional Variation*:** India is a heterogeneous country. Inter-state variation in human development in India is striking, and this not only reflects variations in culture, institutions, and social practises but also variations in the politics, geography, and economics of these states. Gender inequity continues to remain a serious problem in all states, though the degree of heterogeneity is striking – for example, the gender gap in private schooling is only about 7% in Kerala, while it is more than four times that in (30%) in Rajasthan and Bihar. Ceteris paribus, one can attribute this to variations in social practices and exogamy in the North, as opposed to endogamy in the South (see Dyson and Moore (1983)), and also the possible negative effects of endogamy (see Rahman and Rao (2005)), so the net effect is ambiguous.

Rows 25-30 of Table 4 presents the estimated gender gap in female private school enrolment across the regions of the country in 2012. There is clear evidence of significant gender bias (as reflected in the negative and statistically significant GIRL dummy) in all the Indian regions, with the exception of the western states of Gujarat and Maharashtra, which are relatively more industrialised and thus may offer more employment/earning opportunities for girls. Additionally there is a large variation in the coefficient estimate of the GIRL dummy across the various regions: Girls are almost 10 percentage points less likely to attend private school in the Northern states, as compared to about 6 percentage points in North Western and southern regions of the country; the corresponding effect is much smaller for the Eastern (2 percentage points) and the North Eastern (4 percentage points) states. In other words, these results question the simple north-south dichotomy within the context of the country. The tests of equality presented in Column 4 imply that the extent of gender bias is significantly greater in the Northern and North Western regions (relative to that in the Western regions where there is no evidence of gender bias against girls in terms of school choice).

2005 Estimates

The corresponding results for the 2005 data are presented in Columns 5-8 of Table 4. As with the 2012 estimates (shown in Columns 1-4 of Table 4), these 2005 estimates correspond to the regression specification (1') with child-invariant household fixed effects. As before, each row presents the results corresponding to a different sub-sample. Comparing Columns 1 and 5 in Table 4, we see that in general, the size of the female disadvantage was lower in 2005: 4.5 percentage points for the 7-18 year olds and the 7-14 year olds and about 4 percentage points for the 15-18 year olds in our sample. Compared to the full sample of all 7-18 year olds, the female disadvantage was 2.4 percentage points for children who had mothers with more than 10 years of schooling, 4.2 percentage points for children who had fathers with more than 10 years of schooling, 4.8 percentage points for upper-caste Hindu households and 3.6 percentage points for Muslim households. Female disadvantage persisted, even when the average female market wage rate was high and also when the local private school was close. Finally, considering the z-statistics for the equality of the GIRL coefficient between each pair of complementary samples, we can reject the null hypothesis of equality in the following cases: (i) non-eldest girls had significantly higher disadvantages, (ii) girls from Christian households are significantly less disadvantaged relative to those in high-caste Hindu households, (iii) girls with mothers having at least 10 years of schooling (relative to those with less than ten years of schooling) were significantly less disadvantaged, and (iv) girls from northern and north-western (relative to western) regions of the country were significantly worse off in this respect.

A comparison of these 2005 results with those for 2012 seems to highlight a worsening schooling situation in that female disadvantage was generally higher in 2012 and also that the relative advantage of girls in cases (ii) and (iii) disappeared, but the significant disadvantage of non-eldest girls and also any girls residing in the northern and north-western regions persisted. In Section 4.4, we will use the 2005 and 2012 matched panel data to further examine the statistical significance of these effects in 2012 relative to those in 2005.

4.2. Comparison of household FE estimates with/without child-varying trends

Table 5 shows the 2012 and 2005 estimates for private school enrolment Equation (2') with both child-invariant and child-varying household fixed effects. This remains our preferred specification because it minimises any omitted-variable bias arising from both unobserved child-invariant and child-varying within-household factors. Columns 1-4 present the 2012 estimates, while columns 5-8 present those for the 2005 IHDS.

<INSERT TABLE 5 HERE>

The estimates presented in Columns 1 and 5 in Table 5 are very similar to those presented in Columns 1 and 5 of Table 4. A significant female disadvantage persists across most samples and sub-samples that we considered, and the average size of this female disadvantage was around 6.3 percentage points in 2012, as compared to 4.0 percentage points in 2005. A comparison of the size of the GIRL coefficients across the two years (2005 and 2012) and also those across various subsamples shown in Table 5 reveals a very similar picture to what we observed in Table 4 (private school estimates without omitted child-varying factors), thus confirming the robustness of these estimates with the inclusion of omitted within-household child-varying factors.

4.3. Further robustness tests

We examine the robustness of our results with regard to two alternative samples. In the first, we restrict the sample of households to those that exhibit within-household variation in school type choice; in the second, we extend the sample to include all children aged 7–18 residing in the household at the time of the survey, i.e., not restricting the analysis to the sample of children born to the household head.

In the regression results presented in Tables 4 and 5, the estimates are computed over all children within the household, irrespective of whether or not there is variation in school choice within the household. We now restrict the sample to those that exhibit within-household variation in school choice – in this sample, some children attend government schools while others attend private schools even when they are born to same parents. We call this the restricted sample with

within-household variation in school choice across children. A priori, we do not know what generates this variation in school choice among children within a household. Hence, as before, we use the standard regression approach with a set of child-level explanatory variables, including the gender of the child, to identify the factors determining the variation in private school choice among children born to the same parents. As in the full sample, here as well, the size of the gender gap is determined by the coefficient estimate of the GIRL dummy.[23]

Before analysing these subsample results, we determine the extent to which individuals in this subsample are comparable with those in the full sample. The results presented in Table A2 in the Appendix show that both the restricted and full sample are comprised of around 48% females, who are aged, on average, around twelve-and-a-half years with no significant statistical differences in means based on age or gender. However, the restricted sample has a lower proportion of the eldest children (32%) as compared to the full sample (54% and 45% in 2005 and 2012, respectively).[24]

<INSERT TABLE 6 HERE>

Panel A of Table 6 summarises the estimates for this restricted sample for all 7-14 and 15-18 year age groups and also for rural and urban regions. Our estimation methodology accounts for both child-invariant and child varying unobserved fixed effects; i.e., these estimates correspond to Equation (2') with respect to the restricted sub-sample. By design, the sample size is considerably smaller here: There were 1,530 such households in 2012 and 1,326 in 2005. A comparison of these estimates with the full sample estimates, presented in Row 1 of Table 5, highlights the significantly

[23] This is also a model with household fixed effects (both child-invariant and child-varying), which minimizes any potential omitted-variable bias in this sample. This implies that selection bias, which is essentially a reflection of omitted variables bias, would also be minimized in this set up.

[24] We present two tests for the homogeneity of means: (i) a t-statistic that does not take into account the possible correlations for observations of children within a household and (ii) an F-statistic (Koch, et al. (1975)) to test the homogeneity of the two samples, which is equivalent to a test of homogeneity for the row (or column) proportions of a two-way contingency table between the particular individual characteristics (female, age, eldest) and the sample type (discriminating or not). In view of the insignificance of the F-stat for female gender and age at a 5% or lower level, we can conclude that there is no significant difference in age or gender between the two samples – restricted and full. The same cannot be said for the eldest child variable. This holds for both the 2005 and 2012 samples.

larger female disadvantage for these groups of households. In particular, relative to the 2012 full sample estimate of 6.3 percentage points, the size of female disadvantage springs up to about 31 percentage points. In other words, a 7-18-year-old GIRL born to the household head in the restricted (full) sample is about 31 (6) percentage points less likely to be enrolled in a private school as compared to a comparable male child. This difference can be attributed to the fact that the average female disadvantage (as captured by the coefficient estimate of the GIRL dummy) in the full sample is smoothed out by households in which there is no variation in school choice. As before, female disadvantage is significant and persists irrespective of which particular sub-sample one is using, though the size of the disadvantage varies: it is 41.1 percentage points for the 7-14 year olds, 23.2 percentage points for 15-18 year olds, 32.3 percentage points for the rural residents, and 28.9 percentage points for the urban residents in 2012, which is generally compatible with the pattern that we have seen for the full sample in terms of direction (though not in terms of magnitude). A comparison of the size of female disadvantage in 2005 and 2012 restricted-sample households indicates that the size of female disadvantage in private school enrolment was larger in 2012 (relative to 2005): about 31% in 2012, as opposed to 25% in 2005. A similar pattern holds across the various sub-samples that we consider. However, none of the z-statistics are statistically significant, thus ruling out any significant difference in the female disadvantage between 7-14 and 15-18 year olds and also that between rural and urban regions for both survey rounds.

We next examine the extent of gender bias in an extended sample: all children aged 7-18 and residing in the household at the time of the survey (irrespective of variation in school choice); in other words, we now include not only the children born to household heads but also those who were not born to household heads but reside in the same household. The regression results presented in Panel B (Rows 6-10) of Table 6 include both child-invariant and child-varying omitted factors and confirm that the size and significance of the GIRL dummy are qualitatively and quantitatively similar to the full sample estimates presented in Table 5. Furthermore, the binary indicator of whether the child was born to the household head or not (an additional variable that we included in this regression) turns out to be statistically insignificant. Traditionally, in Indian

society, the household head occupies an important position because s/he is responsible for all the household decisions regarding the children and other family members. These results further justify our focus on children born to household heads in this paper, especially because female disadvantage in private school enrolment is driven by parental preferences; without a doubt, the household head is the relevant entity to consider because only s/he has full decision-making authority.

4.4 Comparison of household fixed-effects estimates with/without selectivity correction

Finally, building on Equation (4), we compare the household fixed-effects (with both child-invariant and child-varying) estimates presented in Table 5 with the household fixed-effects estimates with selectivity correction. Essentially, we first estimate a current school enrolment equation (Equation (3)), which allows us to estimate the inverse Mill's ratio (λ_{ij}) for each sample child, treating the data as pooled (i.e., without accounting for household fixed effects). This is then included as an additional argument in determining school choice (Equation 4) using household fixed effects models. These results are then summarised in Table A3 (enrolment estimates) and Table A4 (private school choice estimates) for both the full and restricted samples in 2005 and 2012. Evidently these selectivity-corrected estimates are rather similar to those presented in Tables 5: female disadvantage in private school enrolment is, on average, 6.6 percentage points in 2012 (the corresponding FE estimate is 6.3) as against 5.1 percentage points (the corresponding FE estimate is around 4) in 2005 for the full sample. As before, these estimates vary across the subsamples considered, but they are around five percentage points or less than the corresponding FE estimates as shown in Table 5.

4.5. Matched Household Panel Results

An important feature of the 2012 IHDS data is that it is possible to trace a large fraction of the households that were surveyed in 2005 seven years later, in 2012. This allows us to explore

whether the pattern of female disadvantage that we have seen separately for the 2005 and 2012 samples also holds in this matched panel. As before, we only consider the panel at the household level because we attempt to identify the nature of household preferences regarding private school choice for their children. As before, we consider 7-18 year olds who are born to household heads.

To account for change over time, we include two additional terms: $Wave\ 2$ (= 1 if survey = 2012 IHDS2) and $GIRL \times Wave\ 2$. Recall that the sample is restricted to households that were surveyed in both rounds. The regression specification is as follows:

$$S_{ij}^{*} = \beta_0 + \beta_1 GIRL_i + \beta_2 Wave\ 2 + \beta_3 (GIRL_i \times Wave\ 2) + \beta_4 Eldest_i + \sum_k \alpha_k I(Age = k) + \eta_j + (\theta_i \times \eta_j) + v_{ij} \quad (4')$$

The coefficient estimate of the interaction term (β_3) is equivalent to a difference-in-difference estimate (the difference in the likelihood of girls attending private school across the two waves and the net difference in the likelihood of boys attending private school across the two waves). Recall that the sample is restricted to children aged 7-18 who were born to household heads and residents of households that were surveyed in both rounds. Again, we present estimates not only for the full sample but also for selected sub-samples because understanding the variation in parental son preferences across these sub-samples is also of interest. The results are summarised in Table 7.

<INSERT TABLE 7 HERE>

The GIRL dummy is negative and statistically significant for most sub-samples, except 15-18 year olds, Muslims, Christians, and daughters of mothers with at least 10 years of schooling, which replicates the 2005 estimates for private school enrolment. The coefficient of $Wave\ 2$ is positive and statistically significant for most sub-samples, indicating a higher likelihood of private school enrolment in 2012, which is compatible with known trends. However the interaction term $GIRL_i \times Wave$ is not statistically significant, indicating that there is no significant difference in female disadvantage in 2012 relative to 2005; in other words, girls continued to be disadvantaged in 2012.

5. Concluding Comments

Female disadvantage in schooling, as in other spheres of life, is pronounced in India. Securing universal education for girls remains a priority for the government of India as it strives for inclusive growth. About one-third of school-age children currently attend private schools, though we know very little about female disadvantage in private schooling. Our knowledge of gender inequity in schooling would remain incomplete if we ignored the role of gender in private school choice. The present paper aims to bridge this gap in the literature. We use the 2005 and 2012 India Human Development Survey data to examine the role of gender in private school choice in India.

Using a sample of 7-18-year-old children born to household heads in 2005 and 2012 and minimising the endogeneity of gender, we find that boys are significantly more likely to be enrolled in private schools than comparable girls in the same households. The size of this female disadvantage is significantly higher among non-eldest girls and also girls residing in the northern and north-western (relative to western) regions; in contrast, it is lower in relatively poor households, girls born to educated mothers with at least ten years of schooling, and also to Christian mothers. Also, the gender of a child matters in determining private school choice among restricted households (where school choice varies among children born to same parents), and the size of the female disadvantage here is larger than the full sample. The results hold even when we add children born to non-household-heads residing in the households to the 2005 and 2012 matched household panel. Thus, there is no denying the fact that female disadvantage in private schooling exists in both waves of the IHDS and its size varies across sub-samples and over time, highlighting the non-altruistic nature of preferences: parents tend to balance economic and other socio-cultural considerations (some of which are deep-rooted values) as boys and girls progress from childhood into adolescence.

How does one address the issue of gender bias against girls? Household preferences are often deeply held convictions and difficult to change in the short run (see, for example, Dhar, et al. (2014). Our results highlight the fact that education for all cannot be secured by ignoring the

rapidly growing proportion (about 31% in 2012) of children attending private schools. Rather there must be co-ordinated policies to regulate the growth of public and private schools. In addition to incentivising school participation at all levels for all children, pro-girl policy interventions can go a long way to encourage girls to complete schooling; these would include the adaptation of the school curricula to meet girls' needs, e.g., the recruitment of female teachers; adjusting school hours so as to accommodate household chores for girls; ensuring local access to schools and/or transport to and from schools, especially at the secondary level; and the provision of information about job opportunities for girls. Extending affirmative action policies to ensure girls are enrolled in private schools is a policy worth examining, and incentivizing parents to send their daughters to private schools would help in minimizing this gender gap.

References

ANDRABI, T., J. DAS, A. I. KHWAJA, T. VISHWANATH, & T. ZAJONC (2008): "Learning and Educational Achievements in Punjab Schools: Insights to Inform the Policy Debate," Discussion paper, World Bank.

BECKER, G., & H. G. LEWIS (1965): "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy*, S279 - S288.

BEHRMAN, J. R. (1988): "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favoured? Do Parents Exhibit Inequality Aversion?," *Oxford Economic Papers*, 40, 32 - 54.

BROWNING, M., & P.-A. CHIAPPORI (1998): "Efficient Intra-Household Allocation: A General Characterization and Empirical Tests," *Econometrica*, 66, 1241 - 1278.

BURDE, D., & L. LINDEN (2013): "The Effects of Village-Based Schools: Evidence from a Randomized Controlled Trial in Afghanistan," *American Economic Journal: Applied Economics*, 5, 27 - 40.

CEDAW (2012): "Gender Discrimination in Education: The Violation of Rights of Women and Girls," Global Campaign for Education. Report submitted to the Committee on the Elimination of Discrimination against Women.

CHAUDHURY, N., J. HAMMER, M. KREMER, K. MURALIDHARAN, & F. HALSEY (2006): "Missing in Action," *Journal of Economic Perspectives*, 20, 91 - 116.

CHUDGAR, A., & E. QUIN (2012): "Relationship between Private Schooling and Achievement: Results from Rural and Urban India," *Economics of Education Review*, 31, 376 - 390.

CLOGG, C. C., E. PETKOVA, & A. HARITOU (1995): "Statistical Methods for Comparing Regression Coefficients between Models," *American Journal of Sociology*, 100, 1261 - 1293.

DAS GUPTA, M. (1987): "Selective Discrimination against Female Children in Rural Punjab, India," *Population and Development Review*, 13, 77 - 100.

DAVIES, J. B., & J. ZHANG (1995): "Gender Bias, Investments in Children, and Bequests," *International Economic Review*, 36, 795 - 818.

DHAR, D., S. JAYACHANDRAN, & T. JAIN (2014): "Intergenerational Transmission of Gender Attitudes: Evidence from India," NBER working paper # 21429.

DRÈZE, J., & G. G. KINGDON (2001): "School Participation in Rural India," *Review of Development Economics*, 5, 1 - 24.

DRÈZE, J., & A. SEN (2013): *An Uncertain Glory: India and Its Contradictions*. Oxford University Press, New Delhi.

DURAISAMY, P. (2002): "Changes in the Returns to Education in India 1983-94: By Gender, Age-Cohort and Location," *Economics of Education Review*, 21, 609 - 622.

DYSON, T., & M. MOORE (1983): "On Kinship Structure, Female Autonomy, and Demographic Behavior in India," *Population and Development Review*, 9, 35 - 60.

FILMER, D. (2005): "Gender and Wealth Disparities in Schooling: Evidence from 44 Countries," *International Journal of Educational Research* 43, 351 - 369.

FUJIWARA, T. (2015): "Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil," *Econometrica*, 83, 423 - 464.

GARG, A., & J. MORDUCH (1998): "Sibling Rivalry and the Gender Gap: Evidence from Child Health Outcomes in Ghana," *Journal of Population Economics*, 11, 471 - 493.

GLICK, P. (2008): "What Policies Will Reduce Gender Schooling Gaps in Developing Countries: Evidence and Interpretation," *World Development*, 36, 1623 - 1646.

HÄRMÄ, J. (2009): "Can Choice Promote Education for All? Evidence from Growth in Private Primary Schooling in India," *Compare: A Journal of Comparative and International Education*, 39, 151 - 165.

HEATH, R., AND M. MOBARAK (2011): "Evidence from Garment Sector Jobs and the Female Stipend Program in Bangladesh," Yale School of Management.

HECKMAN, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153 - 161.

JAYACHANDRAN, S., & I. KUZIEMKO (2011): "Why Do Mothers Breastfeed Girls Less Than Boys: Evidence and Implications for Child Health in India," *Quarterly Journal of Economics*, 126, 1485 - 1538.

JENSEN, R. (2012): "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India," *Quarterly Journal of Economics*, 127, 753 - 792.

JHA, P., M. A. KESLER, R. KUMAR, F. RAM, U. RAM, L. ALEKSANDROWICZ, D. G. BASSANI, S. CHANDRA, & J. K. BANTHIA (2011): "Trends in Selective Abortions of Girls in India: Analysis of Nationally Representative Birth Histories from 1990 to 2005 and Census Data from 1991 to 2011," *Lancet*, 377, 1921 - 1928.

KINGDON, G. G. (1996): "The Quality and Efficiency of Public and Private Schools: A Case Study of Urban India," *Oxford Bulletin of Economics and Statistics*, 58, 57 - 82.

— (1998): "Does the Labour Market Explain Lower Female Schooling in India?," *Journal of Development Studies*, 35, 39 - 65.

— (2002): "The Gender Gap in Educational Attainment in India: How Much Can Be Explained?," *Journal of Development Studies*, 39, 25 - 53.

— (2007): "The Progress of School Education in India," *Oxford Review of Economic Policy*, 23, 168 - 195.

KINGDON, G. G., & N. THEOPOLD (2006): "Do Returns to Education Matter to Schooling Participation?," Global Poverty Research Group working paper 52.

KISHOR, S. (1993): "May God Give Sons to All: Gender and Child Mortality in India," *American Sociological Review*, 58, 247 - 265.

KOCH, G. G., D. H. FREEMAN, & J. L. FREEMAN (1975): "Strategies in the Multivariate Analysis of Data from Complex Surveys," *International Statistical Review*, 43, 59 - 78.

KORSON, J. H. (1971): "Endogamous Marriage in a Traditional Muslim Society: West Pakistan: A Study in Intergenerational Change," *Journal of Comparative Family Studies*, 2, 145 - 155.

MAITRA, P., S. PAL, & A. SHARMA (2013): " What Explains the Gender Gap in Private School Enrolment? Evidence from India," University of Surrey and IZA.

MITTERAUER, M. (1991): "Christianity and Endogamy," *Continuity and Change*, 6, 295 - 333.

MURALIDHARAN, K., & M. KREMER (2006): "Public and Private Schools in Rural India," Mimeo Harvard University.

MURTHI, M., A.-C. GUIO, & J. DRÈZE (1995): "Mortality, Fertility, and Gender Bias in India: A District-Level Analysis," *Population and Development Review*, 21, 745 - 782.

NICHOLS, A., & M. SCHAFFER (2007): "Clustered Errors in Stata," Stata Corporation. .

PAL, S. (1999): "An Analysis of Childhood Malnutrition in Rural India: Role of Gender, Income and Other Household Characteristics," *World Development*, 27, 1151 - 1171.

— (2004): "How Much of the Gender Difference in Child School Enrolment Can Be Explained? Evidence from Rural India," *Bulletin of Economic Research*, 56, 133 - 158.

PANDE, R., & A. MALHOTRA (2006): "Son Preference and Daughter's Neglect in India," UNFPA report, International Centre for Research on Women.

PATERNOSTER, R., R. BRAME, P. MAZEROLLE, AND A. PIQUERO (1998): "Using the Correct Statistical Test for Equality of Regression Coefficients," *Criminology*, 36, 859 - 866.

QURESHI, J. A. (2013): "Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital," Discussion paper, University of Illinois and Chicago.

RAHMAN, L., & V. RAO (2005): "The Determinants of Gender Equity in India: Examining Dyson and Moore's Thesis with New Data," *Population and Development Review*, 30.

ROSENZWEIG, M. R., & K. WOLPIN (2000): "Natural Natural Experiments in Economics," *Journal of Economic Literature*, 38, 827 - 874.

SCHULTZ, T. P. (2001): "Why Governments Should Invest More to Educate Girls," Economic Growth Center, Yale University.

SEMYKINA, A., & J. M. WOOLDRIDGE (2013): "Estimation of Dynamic Panel Data Models with Sample Selection," *Journal of Applied Econometrics*, 28, 47 - 61.

SRIVASTAVA, P. (2006): "Private Schooling and Mental Models About Girls' Schooling in India," *Compare: A Journal of Comparative and International Education*, 36, 497 - 514.

SUNDARAM, A., & R. VANNEMAN (2008): "Gender Differentials in Literacy in India: The Intriguing Relationship with Women's Labor Force Participation," *World Development*, 36, 128 - 143.

TOOLEY, J., & P. DIXON (2003): "Private Schools for the Poor: A Case Study from India," Discussion paper, CfBT Research and Development.

UN (2000): "United Nations Millennium Declaration,," Discussion paper, United Nations, New York.

WOODHEAD, M., M. FROST, & Z. JAMES (2013): "Does Growth in Private Schooling Contribute to Education for All? Evidence from a Longitudinal, Two Cohort Study in Andhra Pradesh, India," *International Journal of Educational Development*, 33, 65 - 73.

WOOLDRIDGE, J. M. (1995): "Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions," *Journal of Econometrics*, 68, 115 - 132.

Figure 1: Enrolment Rate by Gender and Survey Wave

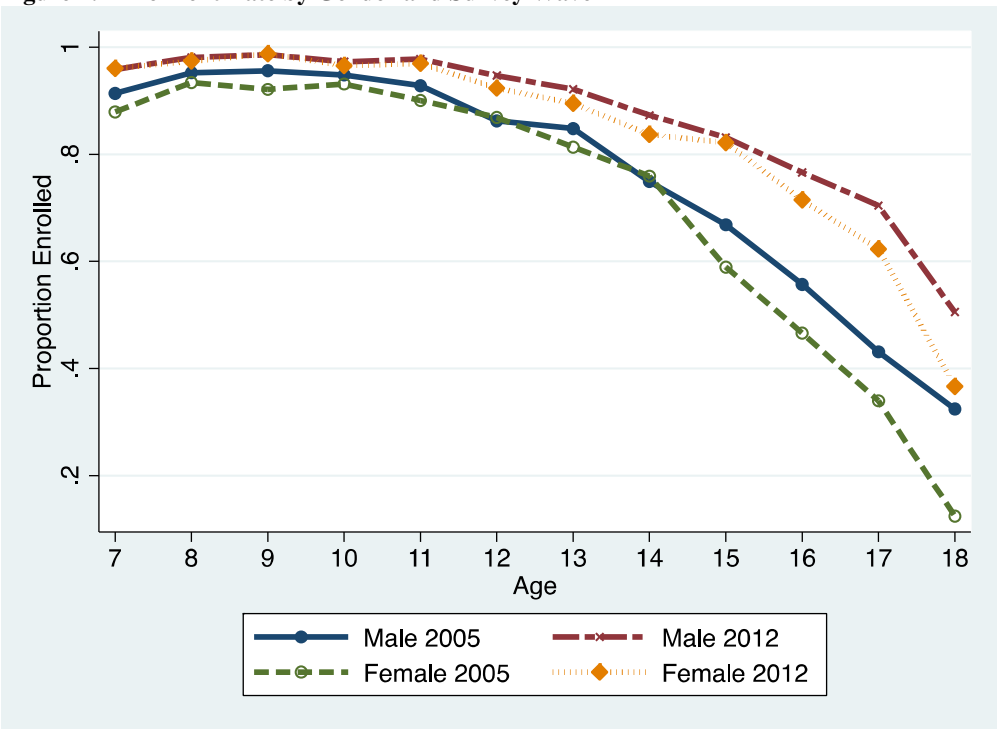
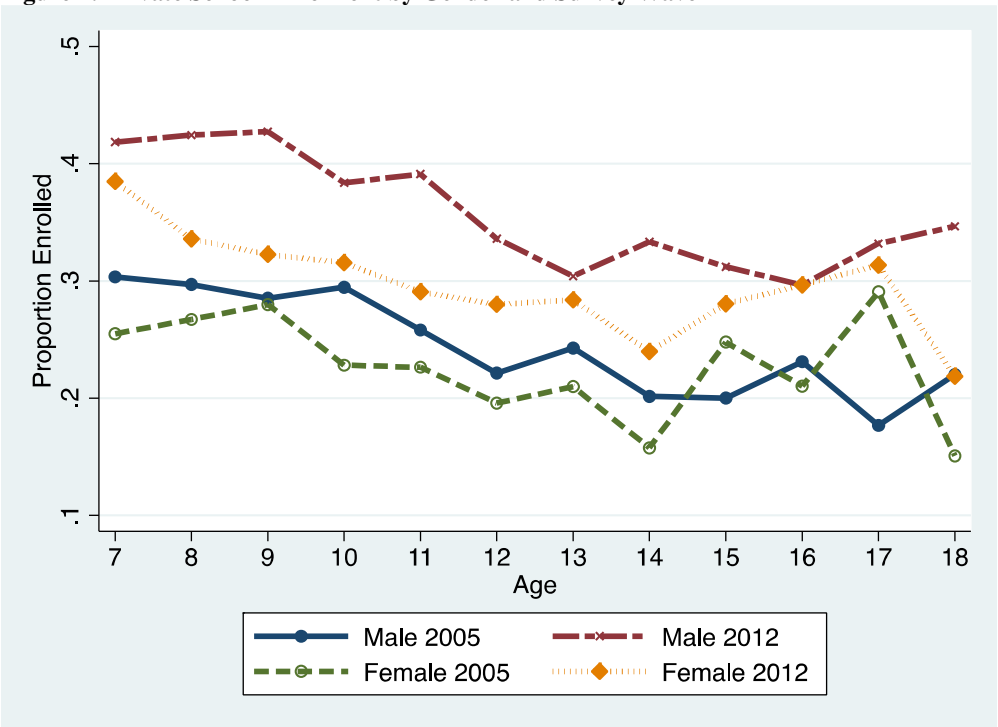


Figure 2: Private School Enrolment by Gender and Survey Wave



Notes: Enrolment in private school conditional on current enrolment

Table 1. A comparison of government, private aided and private unaided schools, 1992 – 2002

	Private schools		Government schools		Private aided schools	
	1992 Mean (SD)	2002 Mean (SD)	1992 Mean (SD)	2002 Mean (SD)	1992 Mean (SD)	2002 Mean (SD)
Share of female teachers in total						
Primary	0.55 (0.26)	0.55 (0.24)	0.35 (0.23)	0.38 (0.15)	0.48 (0.27)	0.49 (0.26)
Upper primary	0.50 (0.27)	0.50 (0.24)	0.30 (0.21)	0.32 (0.17)	0.43 (0.26)	0.42 (0.24)
Secondary	0.45 (0.26)	0.44 (0.22)	0.28 (0.20)	0.24 (0.19)	0.32 (0.23)	0.33 (0.23)
Share of low caste teachers						
Primary	0.09 (0.13)	0.11 (0.13)	0.22 (0.19)	0.24 (0.19)	0.39 (0.30)	0.47 (0.30)
Upper primary	0.08 (0.13)	0.10 (0.12)	0.17 (0.14)	0.23 (0.17)	0.40 (0.28)	0.47 (0.29)
Secondary	0.07 (0.11)	0.08 (0.08	0.15 (0.13)	0.17 (0.12)	0.33 (0.26)	0.42 (0.27)
Share of schools with:						
Pucca building	0.78 (0.17)	0.77 (0.39)	0.66 (0.24)	0.79 (0.34)	0.20 (0.16)	0.76 (0.24)
Lavatory	0.66 (0.23)	0.71 (0.22)	0.33(0.26)	0.41 (0.27)	0.62 (0.25)	0.86 (0.45)
Drinking water	0.84 (0.17)	0.91 (0.13)	0.58 (0.24)	0.78 (0.17)	0.86 (0.17)	0.91 (0.26)
Pupils per teacher						
Primary	30.7 (12.5)	34.3 (31.6)	39.10 (16.2)	67.1 (70.5)	40.77 (14.81)	66.97 (22.9)
Upper Primary	30.8 (11.2)	20.6 (50.0)	31.50 (11.50)	35.3 (58.3)	38.40 (16.07)	39.99 (53.8)
Secondary	29.1 (10.9)	13.7 (23.0)	28.1 (8.4)	29.7 (19.1)	33.85 (11.68)	38.9 (39.08)

Notes: Source: 6th (1992) and 7th (2002) rounds of All India School Education Survey (AISES) Data. Government schools do not include private aided schools. Unfortunately, more recent and comparable AISES data is not yet available (see endnote 7).

Table 2. Private School enrolment, conditional on current enrolment in any school

	2005					2012				
Category	**Sample Size**	**All**	**Male**	**Female**	**Difference**	**Sample Size**	**All**	**Male**	**Female**	**Difference**
All	41400	0.255	0.274	0.233	0.041***	39896	0.323	0.355	0.288	0.067***
Age 7 – 9	11875	0.281	0.302	0.258	0.044***	10898	0.364	0.402	0.324	0.078***
Age 10 – 14	21238	0.251	0.276	0.222	0.054***	19859	0.313	0.349	0.273	0.076***
Age 15 – 18	8287	0.229	0.232	0.225	0.007	9139	0.296	0.313	0.277	0.037***
Rural	27821	0.172	0.191	0.148	0.043***	27165	0.245	0.279	0.208	0.071***
Urban	13579	0.426	0.449	0.401	0.047***	12731	0.490	0.519	0.459	0.060***
Expenditure Q1	9845	0.105	0.121	0.088	0.033***	10677	0.144	0.170	0.119	0.052***
Expenditure Q2	10032	0.171	0.190	0.150	0.039***	10353	0.250	0.273	0.225	0.048***
Expenditure Q3	10434	0.269	0.283	0.253	0.030***	9944	0.379	0.403	0.350	0.053***
Expenditure Q4	11089	0.451	0.472	0.425	0.047***	8922	0.561	0.590	0.526	0.064***
Hindu High Caste	9214	0.351	0.382	0.314	0.068***	7612	0.446	0.481	0.405	0.076***
Hindu SC/ST/OBC	25569	0.205	0.222	0.186	0.036***	24186	0.277	0.308	0.243	0.065***
Muslim	5349	0.271	0.281	0.260	0.020*	5781	0.339	0.372	0.304	0.068***
Christian	551	0.347	0.367	0.329	0.038	939	0.394	0.414	0.375	0.039
Father Education ≥ 10	11217	0.422	0.443	0.398	0.045***	10637	0.505	0.538	0.468	0.070***
Mother Education ≥ 10	5838	0.502	0.508	0.496	0.012	6654	0.591	0.616	0.564	0.052***
North-west	6547	0.347	0.382	0.306	0.076***	5654	0.379	0.414	0.335	0.079***
North	13095	0.301	0.323	0.273	0.051***	14198	0.396	0.446	0.340	0.106***
East	6263	0.163	0.175	0.149	0.026***	6650	0.160	0.181	0.139	0.042***
North-east	1942	0.177	0.184	0.169	0.015	1572	0.258	0.281	0.235	0.047**
West	5289	0.174	0.178	0.170	0.007	4795	0.250	0.255	0.245	0.011
South	8264	0.250	0.263	0.235	0.028***	7027	0.351	0.367	0.334	0.034***

Notes: Sample restricted to children aged 7 – 18 at the time of survey. Difference computed using a two-sided t-test. Private school enrolment conditional on being enrolled in school at the time of survey. *** p<0.01, ** p<0.05, * p<0.1

Table 3. Descriptive statistics

	2005			2012		
	Sample Size	**Mean**	**SD**	**Sample Size**	**Mean**	**SD**
Female	55545	0.482	0.500	46316	0.484	0.500
Ever attend school	55545	0.902	0.297	46316	0.965	0.184
Enrolled	50114	0.826	0.379	44694	0.893	0.310
Enrolled in a private school	41400	0.255	0.436	39896	0.323	0.468
Urban	55545	0.311	0.463	46316	0.310	0.462
North west	55545	0.144	0.351	46316	0.134	0.341
North	55545	0.331	0.470	46316	0.366	0.482
East	55545	0.167	0.373	46316	0.171	0.376
North east	55545	0.046	0.209	46316	0.039	0.193
West	55545	0.123	0.328	46316	0.120	0.326
South	55545	0.190	0.392	46316	0.170	0.376
Household Size	55545	6.553	2.966	46316	6.269	2.668
Household expenditure per capita	55545	759.141	680.966	46296	1760.929	1713.953
Hindu high caste	55545	0.193	0.395	46316	0.176	0.381
Hindu SC/ST/OBC	55545	0.630	0.483	46316	0.610	0.488
Muslim	55545	0.150	0.357	46316	0.158	0.365
Christian	55545	0.011	0.106	46316	0.022	0.147
Father Education ≥ 10	55545	0.219	0.413	46316	0.237	0.425
Mother Education ≥ 10	55545	0.109	0.312	46316	0.145	0.353
Distance Government/Private School Q4	54398	1.518	0.975	45828	1.897	2.349
Fem/Male Wage ≥ 1 (FM_wage_1)	55243	0.716	0.287	46279	0.712	0.220
Total male 7 – 14	55545	0.989	0.916	46316	0.931	0.854
Total female 7 – 14	55545	0.936	0.949	46316	0.883	0.901
Total male 15 – 18	55545	0.426	0.633	46316	0.345	0.564
Total female 15 – 18	55545	0.426	0.640	46316	0.380	0.594

Notes: Sample restricted to all households that have at least one child in the age group 7 – 18 at the time of the survey.

Table 4: Female Disadvantage in Private Schooling, 2012 and 2005. Children born to Household heads.

		2012				2005			
		GIRL	**Sample Size**	**Number of Households**	**Equality of GIRL coeffs (z-statistic)**	**GIRL**	**Sample Size**	**Number of Households**	**Equality of GIRL coeffs (z-statistic)**
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1)	All	-0.063***	30,657	16,921		-0.040***	32,825	17,759	
		(0.005)				(0.004)			
(2)	7-14	-0.060***	22,942	13,910		-0.044***	25,745	15,348	
		(0.005)				(0.004)			
(3)	15-18	-0.044**	7,715	6,663	0.818	-0.021	7,080	6,031	1.337
		(0.020)				(0.017)			
(4)	Rural	-0.061***	20,763	11,157		-0.042***	21,407	11,391	
		(0.006)				(0.005)			
(5)	Urban	-0.066***	9,894	5,764	-0.443	-0.034***	11,418	6,368	0.941
		(0.009)				(0.008)			
(6)	Non Eldest	-0.073***	16,033	10,565		-0.041***	18,982	12,140	
		(0.007)				(0.006)			
(7)	Eldest	0.034	14,624	14,542	14.714***	0.179	13,843	13,765	38.908***
		(0.000)				(0.000)			
(8)	Expenditure Q1	-0.031***	7,270	3,588	-3.317***	-0.030***	7,712	4,019	-0.668
		(0.008)				(0.007)			
(9)	Expenditure Q2	-0.063***	7,504	3,935	-0.810	-0.043***	7,916	4,122	-0.440
		(0.009)				(0.008)			
(10)	Expenditure Q3	-0.083***	7,727	4,292	0.526	-0.042***	8,281	4,446	0.396
		(0.010)				(0.008)			
(11)	Expenditure Q4	-0.075***	8,156	5,106		-0.038***	8,916	5,172	
		(0.011)				(0.009)			
(12)	Hindu High caste	-0.069***	5,196	3,100		-0.046***	6,835	3,857	
		(0.013)				(0.010)			
(13)	Hindu SC/ST/OBC	-0.070***	19,098	10,454	-0.061	-0.040***	20,637	11,149	0.520
		(0.006)				(0.005)			
(14)	Muslim	-0.048***	4,629	2,356	1.192	-0.036***	4,449	2,234	0.680
		(0.012)				(0.011)			
(15)	Christian	-0.035	758	439	0.901	0.051*	444	265	3.126***
		(0.036)				(0.029)			

(16)	Other Religion	-0.020	976	572	1.620	-0.011	460	254	1.021
		(0.027)				(0.032)			
(17)	Father School High	-0.072***	7,698	4,472		-0.034***	8,750	4,841	
		(0.010)				(0.009)			
(18)	Father School Low	-0.060***	22,959	12,450	-0.995	-0.041***	24,075	12,919	0.737
		(0.005)				(0.005)			
(19)	Mother School High	-0.043***	4,559	2,881		-0.014	4,425	2,673	
		(0.013)				(0.013)			
(20)	Mother School Low	-0.066***	26,098	14,047	1.588	-0.043***	28,400	15,092	2.229**
		(0.005)				(0.004)			
(21)	Distance Priv/Gov Q4	-0.056***	7,886	4,337		-0.040***	11,224	6,015	
		(0.008)				(0.006)			
(22)	Distance Priv/Gov Q1 - Q3	-0.066***	22,771	12,584	1.094	-0.040***	21,601	11,744	0.046
		(0.006)				(0.005)			
(23)	Fem/Male Wage >= 1	-0.049***	2,363	1,275		-0.038***	4,956	2,562	
		(0.015)				(0.010)			
(24)	Fem/Male Wage < 1	-0.065***	28,294	15,646	0.962	-0.040***	27,869	15,197	0.121
		(0.005)				(0.004)			
(25)	North West	-0.060***	4,043	2,226	3.550***	-0.058***	4,857	2,425	3.471***
		(0.012)				(0.011)			
(26)	North	-0.107***	10,945	5,524	6.980***	-0.061***	10,211	5,248	4.341***
		(0.009)				(0.008)			
(27)	East	-0.020**	5,321	2,995	1.397	-0.031***	5,073	2,874	1.555
		(0.009)				(0.010)			
(28)	North East	-0.042**	1,383	819	1.983**	-0.003	1,802	947	-0.519
		(0.018)				(0.013)			
(29)	West	0.002	3,414	1,957		-0.011	4,080	2,231	
		(0.013)				(0.008)			
(30)	South	-0.057***	5,551	3,400	3.397***	-0.026***	6,802	4,034	1.325
		(0.012)				(0.009)			
Household Fixed Effect			Yes				Yes		
Within Household Trend			No				No		

Notes: Each row presents the regression on a separate sample. Regressions also control for birth order (whether or not the child is the eldest) and age dummies 8-18 years. Sample includes all children aged 7 – 18 as of the time of the survey and born to the household head. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Columns 4 and 8 present the z-values for the sub-sample tests. Row 3: Age 7-14 vs Age 15-18; Row 5: Rural vs Urban; Row 7: Eldest vs Non-eldest; Row 9: Expenditure Q4 vs Expenditure Q1; Row 10: Expenditure Q4 vs Expenditure Q2; Row 11: Expenditure Q4 vs Expenditure Q3; Row 13: SC/ST/OBC vs Hindu High Caste; Row 14: Muslim vs Hindu High Caste; Row 15: Christian vs Hindu High Caste; Row 16: Other Religion vs Hindu High Caste; Row 18: Father's education High vs Father's education Low;

Row 20: Mother's education High vs Mother's education Low; Row 22: Distance Priv/Gov Q4 vs Distance Priv/Gov Q1 - Q3; Row 24: Fem/Male Wage >= 1 vs Fem/Male Wage < 1; Row 25: North West vs West; Row 26: North vs West; Row 27: East vs West; Row 28: North East vs West; Row 30: South vs West. Test of equality of the GIRL coefficient in two independent subsamples is based on the Z-statistic (see section 3). In doing so, we compare estimates for 15-18 year olds with 7-14 year olds, urban with rural, eldest with non-eldest, SC/ST/OBC, Muslims and Christians with Hindu high caste, each expenditure quartile Q1, Q2 and Q3 with Q4 and each included region with the western region.

Table 5: Female Disadvantage in Private Schooling, 2012 and 2005. Children born to Household heads including within household child trends

		2012				2005			
		GIRL	Sample Size	Number of Households	Equality of GIRL coeffs (z-statistic)	GIRL	Sample Size	Number of Households	Equality of GIRL Coeff (z statistic)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1)	All	-0.063***	30,657	16,921		-0.040***	32,825	17,759	
		(0.005)				(0.004)			
(2)	7-14	-0.060***	22,942	13,910		-0.044***	25,745	15,348	
		(0.005)				(0.004)			
(3)	15-18	-0.046**	7,715	6,663	0.672	-0.023	7,080	6,031	1.153
		(0.020)				(0.017)			
(4)	Rural	-0.061***	20,763	11,157		-0.042***	21,407	11,391	
		(0.006)				(0.005)			
(5)	Urban	-0.066***	9,894	5,764	-0.444	-0.034***	11,418	6,368	0.901
		(0.009)				(0.008)			
(6)	Non Eldest	-0.074***	16,033	10,565		-0.042***	18,982	12,140	
		(0.007)				(0.006)			
(7)	Eldest	0.060	14,624	14,542	18.329***	0.100	13,843	13,765	24.913***
		(0.000)				(0.000)			
(8)	Expenditure Q1	-0.031***	7,270	3,588	-3.371***	-0.029***	7,712	4,019	-1.263
		(0.008)				(0.007)			
(9)	Expenditure Q2	-0.063***	7,504	3,935	-0.866	-0.043***	7,916	4,122	-0.023
		(0.009)				(0.008)			
(10)	Expenditure Q3	-0.082***	7,727	4,292	0.502	-0.043***	8,281	4,446	0.461
		(0.010)				(0.008)			
(11)	Expenditure Q4	-0.074***	8,156	5,106		-0.037***	8,916	5,172	
		(0.011)				(0.009)			
(12)	Hindu High caste	-0.069***	5,196	3,100		-0.045***	6,835	3,857	
		(0.013)				(0.010)			
(13)	Hindu SC/ST/OBC	-0.070***	19,098	10,454	-0.120	-0.040***	20,637	11,149	0.441
		(0.006)				(0.005)			
(14)	Muslim	-0.047***	4,629	2,356	1.261	-0.037***	4,449	2,234	0.526
		(0.012)				(0.011)			
(15)	Christian	-0.035	758	439	0.876	0.054*	444	265	3.173***
		(0.036)				(0.030)			

(16)	Other Religion	-0.024	976	572	1.482	-0.018	460	254	0.782
		(0.028)				(0.032)			
(17)	Father School High	-0.071***	7,698	4,472		-0.034***	8,750	4,841	
		(0.010)				(0.009)			
(18)	Father School Low	-0.060***	22,959	12,450	-0.956	-0.041***	24,075	12,919	0.684
		(0.005)				(0.005)			
(19)	Mother School High	-0.044***	4,559	2,881		-0.013	4,425	2,673	
		(0.013)				(0.012)			
(20)	Mother School Low	-0.066***	26,098	14,047	1.531	-0.043***	28,400	15,092	2.246**
		(0.005)				(0.004)			
(21)	Distance Priv/Gov Q4	-0.056***	7,886	4,337		-0.039***	11,224	6,015	
		(0.008)				(0.006)			
(22)	Distance Priv/Gov Q1 - Q3	-0.066***	22,771	12,584	1.006	-0.040***	21,601	11,744	0.038
		(0.006)				(0.005)			
(23)	Fem/Male Wage >= 1	-0.048***	2,363	1,275		-0.040***	4,956	2,562	
		(0.015)				(0.010)			
(24)	Fem/Male Wage < 1	-0.065***	28,294	15,646	1.048	-0.040***	27,869	15,197	-0.069
		(0.005)				(0.004)			
(25)	North West	-0.059***	4,043	2,226	3.538***	-0.058***	4,857	2,425	3.555***
		(0.012)				(0.011)			
(26)	North	-0.107***	10,945	5,524	6.992***	-0.062***	10,211	5,248	4.479***
		(0.009)				(0.008)			
(27)	East	-0.019**	5,321	2,995	1.410	-0.030***	5,073	2,874	1.522
		(0.009)				(0.010)			
(28)	North East	-0.040**	1,383	819	1.888*	-0.004	1,802	947	-0.344
		(0.018)				(0.012)			
(29)	West	0.003	3,414	1,957		-0.010	4,080	2,231	
		(0.013)				(0.008)			
(30)	South	-0.059***	5,551	3,400	3.508***	-0.026***	6,802	4,034	1.371
		(0.012)				(0.009)			
Household Fixed Effect			Yes				Yes		
Within Household Trend			Yes				Yes		

Notes: Each row presents the regression on a separate sample. Regressions also control for birth order (whether or not the child is the eldest) and age dummies 8-18 years. Sample includes all children aged 7 – 18 as of the time of the survey and born to the household head. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Columns 4 and 8 present the z-values for the sub-sample tests. Row 3: Age 7-14 vs Age 15-18; Row 5: Rural vs Urban; Row 7: Eldest vs Non-eldest; Row 9: Expenditure Q4 vs Expenditure Q1; Row 10: Expenditure Q4 vs Expenditure Q2; Row 11: Expenditure Q4 vs Expenditure Q3; Row 13: SC/ST/OBC vs Hindu High Caste; Row 14: Muslim vs Hindu High Caste; Row 15: Christian vs Hindu High Caste; Row 16: Other Religion vs Hindu High Caste; Row 18: Father's education High vs Father's education Low;

Row 20: Mother's education High vs Mother's education Low; Row 22: Distance Priv/Gov Q4 vs Distance Priv/Gov Q1 - Q3; Row 24: Fem/Male Wage >= 1 vs Fem/Male Wage < 1; Row 25: North West vs West; Row 26: North vs West; Row 27: East vs West; Row 28: North East vs West; Row 30: South vs West. Test of equality of the GIRL coefficient in two independent subsamples is based on the Z-statistic (see section 3). In doing so, we compare estimates for 15-18 year olds with 7-14 year olds, urban with rural, eldest with non-eldest, SC/ST/OBC, Muslims and Christians with Hindu high caste, each expenditure quartile Q1, Q2 and Q3 with Q4 and each included region with the western region.

Table 6: Female Disadvantage in Private Schooling. Robustness Results

		2012				2005			
		GIRL	**Sample Size**	**Number of Households**	**Equality of GIRL coeff (z-statistic)**	**GIRL**	**Sample Size**	**Number of Households**	**Equality of GIRL coeffs (z-statistic)**
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Sample of households with within-household variation in schooling choice (Restricted Sample).									
1	All	-0.308***	4,149	1,530		-0.246***	3,608	1,326	
		(0.023)				(0.025)			
2	7 – 14	-0.411***	1,963	792		-0.361***	1,944	782	
		(0.033)				(0.034)			
3	15 – 18	-0.232**	382	187	1.778	-0.209*	265	124	1.200
		(0.095)				(0.122)			
4	Rural	-0.323***	2,722	988		-0.304***	2,102	775	
		(0.028)				(0.032)			
5	Urban	-0.289***	1,427	542	0.718	-0.159***	1,506	551	2.932
		(0.037)				(0.037)			
Household Fixed Effect				Yes				Yes	
Within Household Trend				Yes				Yes	
Panel B: All children in the household aged 7 – 18 (Extended Sample).									
6	All	-0.061***	39,896	21,657		-0.040***	41,400	21,856	
		(0.004)				(0.004)			
7	7 – 14	-0.060***	30,757	18,383		-0.044***	33,113	19,247	
		(0.005)				(0.004)			
8	15 – 18	-0.040**	9,139	7,797	1.122	-0.023	8,287	6,992	1.153
		(0.017)				(0.017)			
9	Rural	-0.062***	27,165	14,374		-0.042***	27,821	14,341	
		(0.005)				(0.005)			
10	Urban	-0.058***	12,731	7,283	0.345	-0.034***	13,579	7,515	0.901

	(0.008)		(0.008)	
Household Fixed Effect		Yes		Yes
Within Household Trend		Yes		Yes

Notes: Notes: Each row presents the regression on a separate sample. Regressions also control for birth order (whether or not the child is the eldest) and age dummies 8-18 years. In Panel A sample includes children aged 7 – 18 as of the time of the survey, born to the household head and resident of households that exhibit within household variation in school type choice. In Panel B sample includes all children aged 7 – 18 as of the time of the survey. In addition to the individual characteristics GIRL, Eldest and age categories 8-18, this regression also included if the child was born to the household head. While estimates of all other variables were very comparable to those in Table 5, the coefficient estimate of household head dummy remained insignificant. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Test of equality of the GIRL coefficient in two independent subsamples is based on the Z-statistic (see section 3). In doing so, we compare estimates for 15-18 year olds with 7-14 year olds, urban with rural regions.

Table 7: Change in private school enrolment over 2005 – 2012: Results from the Matched Panel of Households

		GIRL	Wave 2012	GIRL × Wave 2012	Constant	Sample Size	Number of households
		(1)	(2)	(3)	(4)	(5)	(6)
(1)	All	-0.055***	0.062***	-0.004	0.269***	42,184	17,453
		(0.006)	(0.008)	(0.008)	(0.006)		
(2)	7-14	-0.060***	0.062***	-0.006	0.215***	32,806	15,848
		(0.006)	(0.009)	(0.009)	(0.012)		
(3)	15-18	-0.037	0.093***	0.014	0.181***	9,378	7,750
		(0.032)	(0.029)	(0.037)	(0.050)		
(4)	Non-Eldest	-0.065***	0.062***	-0.004	0.230***	20,631	11,129
		(0.010)	(0.012)	(0.012)	(0.039)		
(5)	Eldest	-0.036*	0.063***	0.022	0.228***	21,553	15,878
		(0.019)	(0.016)	(0.015)	(0.027)		
(6)	Rural	-0.057***	0.068***	-0.000	0.139***	29,285	12,048
		(0.007)	(0.009)	(0.008)	(0.021)		
(7)	Urban	-0.051***	0.051***	-0.007	0.351***	12,899	5,593
		(0.013)	(0.018)	(0.016)	(0.039)		
(8)	Expenditure Q1	-0.030***	0.063***	-0.011	0.065**	11,784	5,452
		(0.009)	(0.014)	(0.012)	(0.031)		
(9)	Expenditure Q2	-0.055***	0.055**	-0.007	0.174***	11,147	5,740
		(0.012)	(0.022)	(0.015)	(0.037)		
(10)	Expenditure Q3	-0.039***	0.143***	-0.042**	0.151***	10,232	5,528
		(0.013)	(0.025)	(0.018)	(0.043)		
(11)	Expenditure Q4	-0.064***	0.043	-0.007	0.434***	9,021	5,097
		(0.017)	(0.028)	(0.021)	(0.052)		
(12)	Hindu high caste	-0.054***	0.058**	-0.000	0.253***	7,190	3,588
		(0.017)	(0.027)	(0.021)	(0.057)		
(13)	Hindu SC/ST/OBC	-0.055***	0.057***	-0.013	0.197***	26,861	11,622
		(0.007)	(0.010)	(0.009)	(0.022)		
(14)	Muslim	-0.025	0.102***	-0.024	0.148***	6,313	2,348
		(0.015)	(0.020)	(0.020)	(0.050)		
(15)	Christian	-0.109	-0.058	0.082	0.542***	850	445
		(0.072)	(0.093)	(0.083)	(0.196)		
(16)	Other Religion	-0.039	-0.130*	0.049	0.430***	970	512
		(0.051)	(0.079)	(0.059)	(0.144)		
(17)	Father School High	-0.047***	0.097***	-0.008	0.297***	8,222	3,733
		(0.016)	(0.025)	(0.020)	(0.052)		
(18)	Father School Low	-0.057***	0.055***	-0.002	0.179***	33,962	14,314
		(0.006)	(0.009)	(0.008)	(0.020)		
(19)	Mother School High	-0.013	0.095***	-0.023	0.340***	4,273	2,165
		(0.024)	(0.037)	(0.030)	(0.078)		
(20)	Mother School Low	-0.059***	0.060***	-0.000	0.186***	37,911	15,656
		(0.006)	(0.008)	(0.008)	(0.019)		
(21)	Distance Priv/Gov Q4	-0.042***	0.064***	-0.011	0.187***	12,125	6,402
		(0.010)	(0.017)	(0.013)	(0.034)		
(22)	Distance Priv/Gov Q1-Q3	-0.059***	0.068***	-0.003	0.193***	30,059	13,602

		(0.008)	(0.011)	(0.010)	(0.023)		
(23)	Fem/Male Wage >= 1	-0.051***	0.032	0.004	0.244***	4,345	2,290
		(0.018)	(0.038)	(0.026)	(0.061)		
(24)	Fem/Male Wage < 1	-0.052***	0.067***	-0.011	0.202***	37,839	16,276
		(0.006)	(0.009)	(0.008)	(0.020)		
(25)	North West	-0.079***	-0.013	0.025	0.212***	5,884	2,348
		(0.015)	(0.022)	(0.020)	(0.043)		
(26)	North	-0.080***	0.104***	-0.028*	0.232***	14,349	5,295
		(0.012)	(0.015)	(0.014)	(0.034)		
(27)	East	-0.035***	-0.010	0.014	0.147***	7,362	3,130
		(0.012)	(0.015)	(0.014)	(0.037)		
(28)	North East	-0.009	0.100***	-0.024	0.181***	1,803	802
		(0.025)	(0.030)	(0.033)	(0.056)		
(29)	West	-0.005	0.080***	0.011	0.151***	5,060	2,210
		(0.015)	(0.024)	(0.021)	(0.053)		
(30)	South	-0.035**	0.087***	-0.007	0.275***	7,726	3,668
		(0.014)	(0.022)	(0.019)	(0.051)		

Notes: Regressions also control for birth order (whether or not the child is the eldest) and age dummies. Sample includes children aged 7 – 18 as of the time of the survey, born to the household head and born in households that were surveyed in both rounds. Regressions control for household fixed effects. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix

Figure A1: Ever Attended School by Gender and Survey Wave

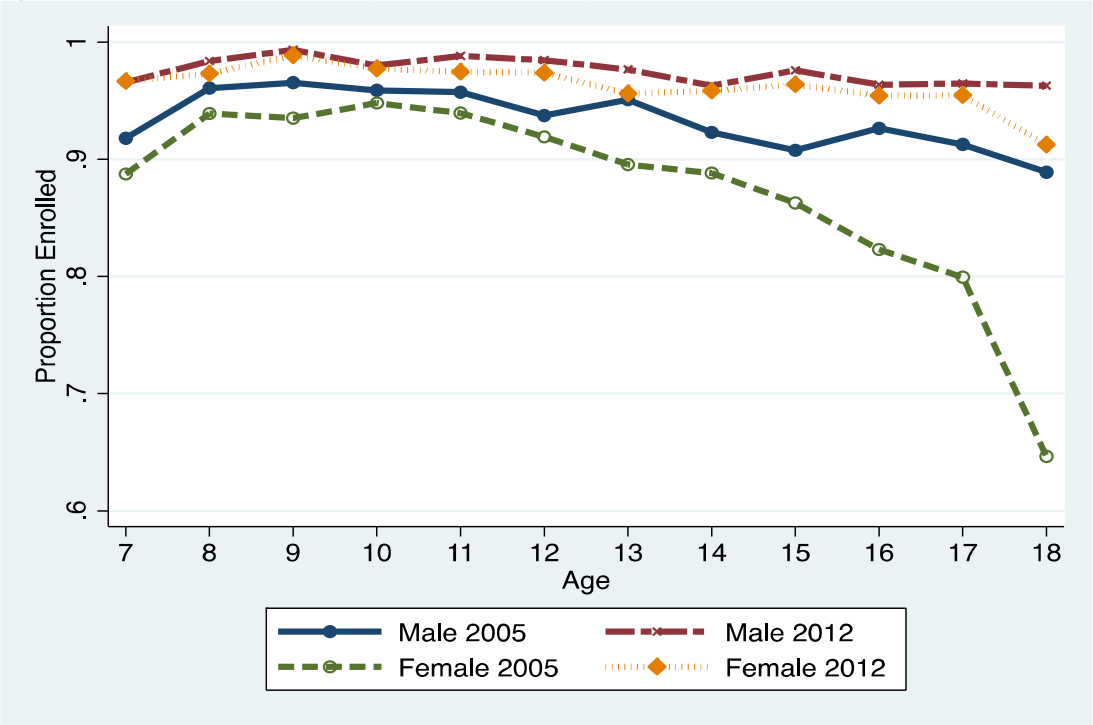


Table A1: Alternatives for 15 – 18 year olds

	Work	Married
Boys, Enrolled in School	0.005	0.006
Boys, Not enrolled in School	0.693	0.034
Girls, Enrolled in School	0.000	0.008
Girls, Not enrolled in School	0.172	0.194

Table A2: Mean comparisons of full and restricted samples with respect to relevant individual characteristics

Variables	Restricted	Non-restricted	T-stat	F-stat (p-value)
2012				
Female	0.487	0.484	0.5812	0.114 (0.7384)
Age	12.43	12.32	2.3487**	2.0053 (0.1563)
eldest	0.32	0.54	31.0823***	142.3 *** (0.00)
2005				
Female	0.476	0.47	0.7463	0.668 (0.4206)
Age	12.3	12.45	2.7185**	3.007 (0.0835)
eldest	0.32	0.45	19.9185***	9.201 *** (0.005)

Note: The table summarises the simple means of the individual level characteristics that is central to estimating the household fixed effects estimates in this paper. We show two tests of homogeneity of means: (i) a t-statistic that does not take account of the possible correlations of observations for children within a household. (ii) An F-statistic (Koch et al. 1975) to test the homogeneity of the two samples, which is equivalent to a test of homogeneity of row (or column) proportions of a two way contingency table between the particular individual characteristic (female, age, eldest) and the sample type (discriminating or not). In view of the insignificance of the F-stat for Female and age at 5% or lower level, we can conclude that there is no significant difference in age or gender in the two samples – restricted and full. The same cannot be said about the eldest variable. This holds for both 2005 and 2012 samples.

Table A3: Probit enrolment estimates from pooled data

	IHDS2 (2012) Enrolment (1)	IHDS1 (2005) Enrolment (2)
GIRL	-0.138***	-0.119***
	(0.028)	(0.018)
Eldest	0.025	0.063***
	(0.037)	(0.021)
Age 7	0.064	0.062
	(0.168)	(0.093)
Age 8	0.168	-0.153*
	(0.194)	(0.087)
Age 9	-0.223	-0.249***
	(0.143)	(0.077)
Age 10	-0.348**	-0.424***
	(0.147)	(0.082)
Age 11	-0.800***	-0.858***
	(0.120)	(0.069)
Age 12	-1.095***	-1.092***
	(0.119)	(0.070)
Age 13	-1.418***	-1.424***
	(0.115)	(0.068)
Age 14	-1.702***	-1.821***
	(0.115)	(0.068)
Age 15	-2.014***	-2.166***
	(0.114)	(0.068)
Age 16	-2.287***	-2.469***
	(0.114)	(0.071)
Age 17	-2.860***	-2.941***
	(0.115)	(0.071)
Urban	0.046	-0.024
	(0.035)	(0.021)
Expenditure Q2	0.112***	0.158***
	(0.038)	(0.025)
Expenditure Q3	0.276***	0.366***
	(0.041)	(0.026)
Expenditure Q4	0.395***	0.578***
	(0.048)	(0.031)
Hindu High caste	0.205**	0.476***
	(0.093)	(0.088)
Hindu SC/ST/OBC	0.001	0.274***
	(0.087)	(0.087)
Muslim	-0.325***	-0.026
	(0.091)	(0.089)
Christian	-0.106	0.534***
	(0.132)	(0.129)
Father Education ≥ 10	0.722***	0.533***
	(0.051)	(0.029)
Mother Education ≥ 10	0.745***	0.609***

	(0.084)	(0.050)
North	-0.280***	-0.198***
	(0.053)	(0.032)
East	-0.152***	-0.332***
	(0.059)	(0.036)
North-east	-0.129	-0.070
	(0.088)	(0.049)
South	-0.161***	-0.128***
	(0.057)	(0.035)
West	-0.450***	-0.300***
	(0.059)	(0.037)
Media regular	0.196***	0.088***
	(0.031)	(0.018)
Constant	2.573***	1.907***
	(0.144)	(0.110)
Sample Size	22,852	40,092
Pseudo R-square	0.3222	0.3376
Wald Chi-square	4402.56***	7822.44***

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A4. Private School Enrolment with selection correction

	All (1)	7-14 (2)	15-18 (3)	Rural (4)	Urban (5)
Panel A: Full Sample 2012					
GIRL	-0.066***	-0.064***	-0.051***	-0.063***	-0.068***
	(0.004)	(0.005)	(0.015)	(0.005)	(0.009)
Constant	0.338***	0.339***	0.281***	0.243***	0.553***
	(0.008)	(0.006)	(0.030)	(0.008)	(0.016)
Household Fixed Effects	Yes	Yes	Yes	Yes	Yes
λ	Yes	Yes	Yes	Yes	Yes
Sample Size	34,626	23,765	10,861	23,666	10,960
Number of households	18,180	14,208	8,938	12,081	6,099
Panel B: Full Sample 2005					
GIRL	-0.051***	-0.048***	-0.062**	-0.055***	-0.037***
	(0.005)	(0.005)	(0.027)	(0.006)	(0.010)
Constant	0.295***	0.283***	0.028	0.190***	0.484***
	(0.007)	(0.007)	(0.099)	(0.00)	(0.0149)
Household Fixed Effects	Yes	Yes	Yes	Yes	Yes
λ	Yes	Yes	Yes	Yes	Yes
Sample Size	21,999	17,984	4,015	14,186	7,813
Number of households	9,457	9,144	3,209	6,039	3,418
Panel C: Restricted Sample 2012					
GIRL	-0.283***	-0.400***	-0.230***	-0.300***	-0.242***
	(0.019)	(0.031)	(0.070)	(0.024)	(0.032)
Constant	0.756***	0.771***	1.106**	0.680***	0.940***
	(0.039)	(0.048)	(0.447)	(0.046)	(0.071)
Household Fixed Effects	Yes	Yes	Yes	Yes	Yes
λ	Yes	Yes	Yes	Yes	Yes
Sample Size	5,155	2,146	627	3,343	1,812
Number of households	1,827	856	306	1,171	656
Panel D: Restricted Sample 2005					
GIRL	-0.268***	-0.365***	-0.221*	-0.356***	-0.152***
	(0.0213)	(0.0313)	(0.116)	(0.0367)	(0.0368)
Constant	0.708***	0.667***	0.0725	0.677***	0.746***
	(0.0298)	(0.0332)	(0.386)	(0.0411)	(0.0501)
Household Fixed Effects	Yes	Yes	Yes	Yes	Yes
λ	Yes	Yes	Yes	Yes	Yes
Sample Size	3,194	1,671	261	1,751	1,443
Number of households	1,116	663	122	621	495

Notes: Sample includes children aged 7 – 18 as of the time of the survey and born to the household head. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1