Intra-Household Gender Discrimination in School Choice: Evidence from Private Schooling in India

Soham Sahoo*

Preliminary draft; March, 2014

Abstract

This paper explores the incidence of gender discrimination within households in the decision of private versus government school choice in India. Recently, there has been a rapid rise in the number of private schools operating in rural areas. Households perceive these fee-charging private schools to be of better quality than government schools which are mostly free. If the future returns from investing in sons' education seem to be higher than daughters' education, then households may want to provide their sons, rather than daughters, with better quality education by sending them to private schools. Using a three-period longitudinal data on rural households, this paper estimates a correlated unobserved effects model with selection correction and finds that households indeed discriminate against girls in favour of boys for private school enrollment. The gender gap in private school enrollment is around six percentage points. This gap is higher among younger children and is rising over time. This finding indicates that the quality of government schools may have a role to play in promoting gender parity in education.

Keywords: School enrollment, private schooling, gender, longitudinal data, sample selection

JEL codes: I20, I24, I25

*Doctoral Candidate, Economics and Planning Unit, Indian Statistical Institute, Delhi. Email: soham9r@isid.ac.in

1 Introduction

Gender equality is one of the central issues in the discourse of development economics. Various reports recently published by the United Nations and World Bank focus on gender parity, particularly on providing equal opportunity to men and women in the economic sphere [1]. It has been a long lasting agenda on the developing economies and several policies have been debated over provision of equity in education. Gender parity is one of the six goals of the global "Education for All" campaign led by the United Nations Educational, Scientific and Cultural Organization (UNESCO). In India, some of the major public policy initiatives like *Sarva Shiksha Abhiyan* (Education for All) have strived to universalize elementary education and reduce disparity across gender, regions and social-groups. While the government has concentrated on providing free education and improving enrollment at the elementary levels, the quality of education has remained a concern [2]. During the last decade, India has experienced a huge surge in schools established by the private providers. Parents who are not satisfied with the quality of government schools have perceived private schools as a better alternative. In the literature, there is no consensus about the effectiveness of private schools in imparting good quality education. Several studies opine that private schools provide better quality of teaching, teacher absenteeism is less and students' learning outcomes in private schools are higher than that in government schools (Kingdon, 1996; Muralidharan and Kremer, 2008; Tooley et al, 2007). On the other hand, there are studies which find that private schooling has serious equity issues: children from poorer households,

[1] The World Development Report (2012) published by the World Bank is on gender equality and development. The United Nations Educational, Scientific and Cultural Organization (UNESCO) released the World Atlas of Gender Equality in Education (2012) which illustrates the educational trends of girls and boys and how gender disparity has changed over time.

[2] Annual Status of Education Reports (ASER) from 2005 to 2013 show that enrollment rate among the children in the primary school going age group has improved substantially and remained steady at above 95 percent. In contrast, learning outcomes in reading and mathematics have been quite unsatisfactory and without any improvement over time.

lower socio-economic backgrounds, from rural areas and girls are less likely to attend private schools (Harma, 2011; Maitra et al, 2011; Woodhead et al, 2013). Thus, there is an ongoing debate on whether private schools are capable of contributing in the path towards the Millennium Development Goals of universalization of primary education and elimination of gender disparity in primary and secondary education by 2015. In this context, this paper seeks to address the issue of gender discrimination in private school choice within households: whether households choose to discriminate between their sons and daughters while providing them an education which they perceive to be of better quality.

Economic theory predicts that if returns to education are higher for boys than girls, or girls are less likely to participate in the labour market, then under credit constraints, households may find it optimal to invest more on their son's education than their daughter's education (Garg and Morduch, 1998; Leclercq, 2001). In India, cost of participating in the government school is quite low while the private schools charge fees and hence they are more expensive. In the era of rapid growth in private schooling, even the poor households want their children to attend a private school because they perceive it to be of higher quality. However, since these schools are more expensive than government schools, given budget constraint, a rational household would strategically decide about which child to send to private school, because it is not possible to bear the expenses of all the children in private school. In the patrilocal residence system, returns from investment in son's education seems to be higher for such households. Therefore, their optimal behaviour would be to send their sons, rather than daughters, to private schools. However, in a social system dominated by patriarchal values, the intrinsic preference of the households can be so biased against female education that even if the budget constraint is relaxed, they may still send only their sons to private schools and not their daughters.

In this paper, I use a household level panel data from rural India to investigate whether there is intra-household gender gap in private school choice. While a number of studies have extensively analyzed the issue of gender disparity in overall enrollment, grade progression and education expenditure in developing countries (Das Gupta, 1987; Deolalikar, 1993; Sawada and Lokshin, 2009; Azam and Kingdon, 2013), very few papers refer to the boy-bias in private school choice (Maitra et al, 2011; Woodhead et al, 2013). Moreover, most of these studies on private school choice are based on cross-section data and they do not look into the intra-household decision making process. They essentially compare different households to find out whether girls' participation in private schools is lower than boys' participation. However, to address the issue of intra-household gender discrimination, it is imperative to control for unobserved household heterogeneity that may be correlated with the observed explanatory variables including the gender of a child. In addition, the choice of private or government school for a child is observed only if the child is enrolled in school. A regression analysis based on only the enrolled children may result in selection bias in the estimates. Therefore, in this paper, I estimate a variant of a model proposed by Wooldridge (1995) where selection corrected estimates are obtained in a panel data set up, and unobserved household heterogeneity is allowed to be correlated with the independent variables.

The results indicate the presence of a strong gender bias against girl children within the household in the decision of enrolling a child in private school. There is a significant gender gap around six percentage points in the probability of private school enrollment. When I decompose this effect, I find that the practice of gender discrimination is actually rising over time: while the gender gap in private school choice was insignificant in 1997-98, it has become almost 10 percentage points in 2010-11. If we look at the proportion of children going to private schools in recent years, this gender gap would imply that girls are around 20 percent less likely to be sent to a private school. I also find that

the gender gap is higher for younger children who belong to the primary school going age group. For the primary level of education, quality of government primary schools in the village has a significant negative effect on enrollment in private primary schools. Therefore, from a policy perspective, unregulated private schools seem unlikely to be instrumental in narrowing down the gender gap in education.

The rest of the paper is organized as follows. Section 2 and 3 describe the dataset and the empirical methodology respectively. Section 4 discusses the results and section 5 concludes.

2 Data

To investigate the issue of gender discrimination in private school choice, I analyze data from a longitudinal follow up of households first surveyed as a part of the World Bank's Living Standards Measurement Study (LSMS) in Uttar Pradesh, a state in India (this survey is also called the Survey of Living Conditions, or SLC). This is a three period panel data on rural households in 43 villages from eastern and southern Uttar Pradesh[3]. The baseline data was collected in 1997-98 under LSMS. The same set of households were resurveyed in 2007-08 and again in 2010-11 [4] [5]. The survey comprised a village questionnaire and a household questionnaire which contained detailed information on the demographics of each household member and schooling information for every child belonging to the age group of 6 to 19 years.

For the purpose of our study, I concentrate on children who are in the school going age, 6-19 years, at each time point observed in the data. The objective is to examine

[3] Uttar Pradesh (UP) is considered to be one of the most backward states in India. Our sample includes 10 districts in UP. The number of villages in each district varies from 3 to 6 per district.

[4] The second round of survey in 2007-08 was funded by the University of Oxford and the World Bank. The third round in 2010-11 was funded by the Planning and Policy Research Unit of the Indian Statistical Institute, Delhi.

[5] All three surveys were conducted during the same time of the year - from December to April.

the pattern of private school enrollment based on the gender of the child. I estimate the overall enrollment rates for boys and girls separately, and see how they have changed over time. Then I analyze the trend in private school enrollment in a similar manner. Using a panel data allows not only to inspect if there is any gender differential treatment in the decision of private school enrollment, but also to observe any temporal change in this discriminatory behaviour by the households.

According to the estimates from our panel data, there is a convergence in the overall enrollment rates of boys and girls over time (*Figure* 1). In 1997-98, 69 percent of the boys in 6-19 age group were enrolled in school, while among girls, only 50 percent were enrolled at that time. However, this gender gap reduced significantly over time. In 2007-08, 65 percent of the girls as against 69 percent of the boys were enrolled. Finally in 2010-11, 71 percent of the girls, as compared to 72 percent of boys, were enrolled. A t-test for mean comparison suggests that there is no statistically significant difference between the enrollment rates of the boys and girls in 2010-11. On the other hand, if we look at the trends in private schooling for children who are enrolled, there is no such convergence across gender. The enrollment rates for both boys and girls have grown steadily over the period of study, from 24 (24) percent in 1998-98 to 57 (47) percent in 2010-11 for boys (girls) (*Figure* 2). Thus the gender gap measured by the difference between enrollment rates has actually increased over time. Combined with the trends in overall enrollment rates, this suggests that the steep rise in private schooling over the years can be attributed to boys than girls. Moreover, while the gender gap in overall school enrollment tends to disappear over time, the gap in terms of private school participation becomes starker.

3 Methodology

From the descriptive statistics it is revealed that there is a gender gap in the enrollment rates in private school. But this finding is based on just the crude measure of average enrollment rates, hence it is not sufficient to suggest that there is a discriminatory behaviour towards girl's schooling within households. If the gender of a child covaries with other factors that affect the likelihood of enrollment into private schools, then the observed gender difference can be confounded by those factors. Therefore, I set up a multivariate regression model to control for these factors and investigate if households actually discriminate against girls while deciding about the quality of their children's education. This analysis is based on the children in the school going age group of 6 to 19 years.

Consider P^* to be the latent decision making process by the household to enrol a child in a private school. However, we can observe only the binary outcome, P, of this decision, that indicates whether the child goes to a private school ($P = 1$) or a government school ($P = 0$).

$$P^*_{cht} = \mathbf{X}_{cht}\boldsymbol{\beta} + \alpha_h + \varepsilon_{cht} \; ; \; P_{cht} = 1[P^*_{cht} > 0]; \tag{1}$$

$$c = 1, ..., C(h,t); \; h = 1, ..., H(t); \; t = 1, 2, 3.$$

The choice of private versus government school is modeled by *Equation* 1. The subscript c refers to a child in household h at time period t. In any time period t, there are $H(t)$ households who have at least one child in the school going age group and are included in the sample; and there are $C(h,t)$ children in h-th household in t-th time period. $\mathbf{X}$ denotes the vector of explanatory variables that could affect the private school choice, and $\boldsymbol{\beta}$ is the corresponding coefficient vector. $\mathbf{X}$ contains our main variable of interest, the gender of the child, captured by a dummy variable indicating whether the

child is female. It also includes other child, household and region specific variables. The child specific variables are the age of the child, square of age (for possible non-linearity in the effect of age), birth order of the child within the household, dummy variables representing whether the father and mother are literate. The household specific variables that are included are dummy variables reflecting whether the head is literate and whether the head is female, total number of female and male children in the household, household size, and household wealth captured by landholding. Among the region level variables, we have the proportion of private schools among all the schools in the village. Quality of education in the government schools in the village can be an important factor to determine participation in private schools (Alderman et al, 2001; Desai et al, 2008; Pal, 2010). Therefore, I include an index constructed by principal component analysis of different indicators of quality in the village based government primary schools [6] [7]. To reflect the secular rise in private school enrollment over time, year dummies are included. Furthermore, the vector of explanatory variables also contains district-by-time fixed effects to allow for differential rate of growth in private schooling across districts.

In addition to the explanatory variables described above, the model also includes household specific unobserved heterogeneity by the term α_h. This is particularly important given that our objective is to investigate if households, while deciding about enrolling their children in a private school, give more preference to boys than girls. Since we follow the same households over time, and there are multiple children of different gender in a household, it is possible for us to identify the coefficient of female dummy even after controlling for household specific fixed effects. By including these

[6] The features of school quality considered in the analysis are type of structure, main flooring material, whether the school has classrooms, number of classrooms, whether the classes are held inside classrooms, whether the school has usable blackboards, whether desks are provided to the students, whether mid-day meal is provided and the proportion of teachers present on the day of survey.

[7] For those villages where more than one government primary schools are present, I consider the representative school to be the one which has the maximum number of students.

household level fixed effects, we control for unobservable factors that are particular to each household and do not change over time. It also ensures that we focus on the decision making process within household rather than comparing outcomes across different households. It may be the case that households which have more female children, in comparison with their counterpart, prefer government schools to private schools. Therefore if we compare across households, we will observe a gender gap in private schooling not because of discrimination, but because of the difference in the intrinsic preference for these two types of schools. Inclusion of α_h in the model ensures that the findings are not confounded by such unobservable heterogeneity. Moreover, female children tend to end up in larger families because fertility decisions are endogenous and parents prefer to have at least one boy child. Since the number of children is determined endogenously within the household, it is also important to control for these fertility preferences by including household fixed effects (Azam and Kingdon, 2013).

We can estimate *Equation* 1 following a Linear Probability Model and obtain Ordinary Least Square (OLS) estimates of $\boldsymbol{\beta}$. In this method, we can control for the household level fixed effects (α_h) either by taking a time-demeaned transformation of *Equation* 1, or by explicitly including household specific dummy variables. However, note that the choice of school type is observed only for those children who are enrolled. If the decisions of school choice and enrollment are correlated, then estimating the school choice equation considering only the selected sample of enrolled children may lead to biased and inconsistent estimates. This is similar to the standard sample selection problem (Heckman, 1979). Therefore, the decision to enrol is taken into account in the econometric model by the latent variable S^*, and its observable counterpart is caputured by the binary enrollment outcome S, which is one if the child is enrolled in some school, and zero otherwise.

$$S^*_{cht} = \mathbf{Z}_{cht}\boldsymbol{\gamma} + \delta_h + u_{cht} \; ; \; S_{cht} = 1[S^*_{cht} > 0]. \tag{2}$$

The enrollment decision modeled in *Equation* 2 has a form similar to the school choice model. The set of explanatory variables, $\mathbf{Z}$, contains all the variables that are present in $\mathbf{X}$; but for ease of identification, it also contains an additional variable that is validly excluded from *Equation* 1. After controlling for the composition of private and government schools in the village, it is plausible to think that total number of schools in the village will affect only the enrollment decision, but not the private school choice decision. Therefore, I use the total number of all kinds of schools in the village as the variable which is excluded from the main equation, but is included in the selection equation. Unobserved household heterogeneity are considered by the term δ_h. The private school choice variable P_{cht} in *Equation* 1 is observable only when $S_{cht} = 1$.

When the selection process is non-random, the most widely used method in the literature for correcting sample selection bias is the model proposed by Heckman (1979), also known as the "Heckit" model. However, in the presence of unobserved heterogeneity, implementation of Heckit model becomes problematic. While we can estimate the main relationship (*Equation* 1) as a linear model based on the selected sample, the Heckit model requires that we estimate the selection equation (*Equation* 2) by using a probit model. But, probit being a non-linear model, it is not possible to eliminate the fixed effects by taking any within-transformation of the equation. Besides, since probit model employs maximum likelihood estimation, if we attempt to estimate the selection equation including household specific dummy variables to capture unobserved heterogeneity, we will face the "incidental parameters problem" (Neyman and Scott, 1948; Lancaster, 2000). This will lead to inconsistency in the estimates of not only δ_h, but also $\boldsymbol{\gamma}$. On the other hand, failure to account for the unobserved heterogeneity, which may be correlated

with other regressors in the model, may result in biased and incosistent estimates of the parameters of interest. Thus, standard Heckit model is infeasible to solve the selection problem in our context.

Wooldridge (1995) offers a method for correcting for sample selection bias in linear panel data models where unobserved heterogeneity is allowed to be correlated with the observable explanatory variables in both the selection equation and the equation of interest. While this method is conceptually similar to Heckman (1979), it is appropriate for panel data models such as in this paper [8].

Since the choice of private versus government school is observed only for the sample of enrolled children, a sufficient condition for obtaining consistent estimates of $\boldsymbol{\beta}$ by running a pooled OLS model on *Equation* 1 is given by:

$$E(\alpha_h + \varepsilon_{cht} \mid \mathbf{X}_{cht}, S_{cht} = 1) = E(\alpha_h \mid \mathbf{X}_{cht}, S_{cht} = 1) + E(\varepsilon_{cht} \mid \mathbf{X}_{cht}, S_{cht} = 1) = 0.$$

The conditional expectation specified above will not be zero if household heterogeneity is correlated with the explanatory variables or the selection process is non-random. One way to tackle this problem would be to parameterise these conditional expectations and add them to the main equation (Wooldridge, 1995; Dustmann and Rochina-Barrachina, 2007). To derive this estimator in the context of our analysis, based on Wooldridge (1995), I assume the following structure of our econometric model:

(i) To allow the unobserved household heterogeneity in the selection equation (*Equation* 2) to be correlated with the explanatory variables, following Mundlak (1978), I assume that δ_h is a linear function of the within-household average of $\mathbf{Z}_{cht}$ over all children and all time periods. Thus,

$$\delta_h = \eta_0 + \overline{\mathbf{Z}}_h \boldsymbol{\eta} + e_h, \tag{3}$$

[8] Dustmann and Rochina-Barrachina (2007) use this method in a similar set up to estimate the females' wage equations.

where e_h is a random variable independent of other factors, and $\overline{\mathbf{Z}}_h = \frac{1}{T(h)} \sum_t (\frac{1}{C(h,t)} \sum_c \mathbf{Z}_{cht})$ is the household specific average values of the observed explanatory variables, with $C(h,t)$ being the number of children present in h-th household at t-th time period and $T(h) \in \{1,2,3\}$ being the number of time-periods when h-th household has at least one child in the relevant age group so that it is included in the sample of our analysis [9].

(ii) Following (i), the reduced form of the selection equation becomes:

$$S^*_{cht} = \eta_0 + \overline{\mathbf{Z}}_h \boldsymbol{\eta} + \mathbf{Z}_{cht} \boldsymbol{\gamma} + v_{cht}, S_{cht} = 1[S^*_{cht} > 0], \tag{4}$$

where $v_{cht} = e_h + u_{cht}$. Assume that v_{cht} is independent of $\mathbf{Z}_{ch}$, and $v_{cht} \sim Normal(0, \sigma^2_t)$, where $\mathbf{Z}_{ch} = (\mathbf{Z}_{ch1}, ..., \mathbf{Z}_{chT})$.

(iii) In the main equation, let us assume that the household specific unobserved effects has a linear relationship with the household level averages of the explanatory variables (Mundlak, 1978). Hence, I assume that the conditional expectation of α_h given $\mathbf{Z}_{ch}$ and v_{cht} is linear. Thus, we have the following relationship:

$$E(\alpha_h \mid \mathbf{Z}_{ch}, v_{cht}) = \psi_0 + \overline{\mathbf{X}}_h \boldsymbol{\psi} + \pi_t v_{cht}, \tag{5}$$

where $\overline{\mathbf{X}}_h = \frac{1}{T(h)} \sum_t (\frac{1}{C(h,t)} \sum_c \mathbf{X}_{cht})$. Note that under the exclusion restriction, the elements of $\mathbf{Z}$ which are not in $\mathbf{X}$ are independent of α_h and ε_{cht}, hence they do not appear in the above relationship.

(iv) Finally, assume that ε_{cht} is mean independent of $\mathbf{Z}_{ch}$ conditional on v_{cht}, and its conditional expectation is linear in v_{cht}:

[9] Another way of allowing for correlation between the unobserved household specific effects and the explanatory variables is similar to Chamberlain's Correlated Random Effects Model (Chamberlain, 1980; Wooldridge, 2002). This model would assume that δ_h is a linear function of the leads and lags of the explanatory variables. But our data is unbalanced in nature because not all households have children in the school-going age group in all time periods, hence Chamberlain's (1980) specification is not suitable here.

$$E(\varepsilon_{cht} \mid \mathbf{Z}_{ch}, v_{cht}) = E(\varepsilon_{cht} \mid v_{cht}) = \rho_t v_{cht}. \tag{6}$$

Note that we do not observe v_{cht}, rather only the binary enrollment decision (S_{cht}) for each child. Since S_{cht} is a function of $\mathbf{Z}_{ch}$ and v_{cht}, we can apply the law of iterated expectations to *Equation* 5 and 6, and combine them to get the following relation:

$$\begin{aligned} E(\alpha_h + \varepsilon_{cht} \mid \mathbf{Z}_{ch}, S_{cht} &= 1) = \psi_0 + \overline{\mathbf{X}}_h \boldsymbol{\psi} + (\pi_t + \rho_t) E(v_{cht} \mid \mathbf{Z}_{ch}, S_{cht} = 1) \\ &= \psi_0 + \overline{\mathbf{X}}_h \boldsymbol{\psi} + \zeta_t \lambda_{cht}, \end{aligned} \tag{7}$$

where $\zeta_t = \pi_t + \rho_t$, $\lambda_{cht} = E(v_{cht} \mid \mathbf{Z}_{ch}, S_{cht} = 1)$. Finally, the main equation capturing private school choice is modified in accordance with the econometric structure above. Thus, we have:

$$E(P_{cht} \mid \mathbf{X}_{ch}, S_{cht} = 1) = \psi_0 + \overline{\mathbf{X}}_h \boldsymbol{\psi} + \mathbf{X}_{cht} \boldsymbol{\beta} + \zeta_t \lambda_{cht}. \tag{8}$$

A consistent estimate of $\boldsymbol{\beta}$ can be obtained through *Equation* 8 following a few steps. First, the reduced form sample selection equation (*Equation* 4) is estimated using standard probit model separately for each time period, and λ_{cht} are estimated as the ratio of normal density to cumulative distribution function (also known as the Inverse Mills Ratio). Thus, we have separate estimates of λ_{cht}, or the Inverse Mills Ratio, for different time periods. In the next step, these Inverse Mills Ratios are included in the main regression (*Equation* 8) as additional regressors. Then *Equation* 8 is estimated by pooled OLS method. The standard errors are bootstrapped and clustered at the household level.

It is noteworthy that this method allows for possible correlation between the un-

observed household heterogeneity and the observed explanatory variables through the Mundlak formulation. For the regressors which vary within household or over time, the corresponding elements of $\boldsymbol{\beta}$ are identified in this method. On the contrary, due to the Mundlak formulation, it is not possible to separately identify the elements of $\boldsymbol{\beta}$ from the elements of $\boldsymbol{\psi}$ for those regressors which neither vary within household nor over time. Nevertheless, this does not hinder us from identifying the effect of gender. Once the estimation is carried out, we can investigate if it is important to control for unobserved heterogeneity by performing a Wald test for the joint significance of the elements in $\boldsymbol{\psi}$. Similarly, if the null hypothesis of $\boldsymbol{\zeta} = 0$ is rejected, then it would imply that it is crucial to correct for sample selection bias in the regression.

4 Results

4.1 Main Result

Table 1 contains the results from our main regression for all children in the school going age group of 6 to 19 years. The first column shows results under the model of selection correction, while the second column estimates the regression model without taking into account the sample selection problem (i.e., excluding the λ_{cht} terms). It is found that on an average, being a female child reduces the probability of enrollment in a private school by around 6 percentage points. Considering the average enrollment rate of 25 percent in private schools, this effect would imply that girls are almost 20 percent less likely to be enrolled in private schools than boys. From the regression output in the second column, it is found that if we ignore the potential selection bias, then the coefficient is underestimated. Wald test shows that the correlated unobserved effects are highly significant jointly in the regression. Also, the selection effects are jointly significant,

indicating the importance of taking into account the selection bias.

Among the other child level variables, age and birth order have significant effect on private school enrollment. Probability of enrollment in private school increases with age, but the rate of increase reduces with age. There is a strong negative and significant birth order effect. It suggests that parents tend to invest more on education of the first born children and send them to fee-charging private schools, but under resource constraints, probably they are unable to further keep up the investment on the later born children's education. Presence of more private schools in the village also positively affects the likelihood of private school enrollment.

In the next step, I estimate this model separately for children who are in the primary school going age group and for those who are relatively older and belong to the post-primary (middle, secondary or senior secondary) school going age group. *Table* 2 shows that the gender gap is quite higher for primary level as compared to post primary level. In primary school going age group, girls are around 9 percentage points less likely to be enrolled in a private school vis-a-vis government school, while this gender gap is lesser at 4.3 percent in the post-primary level. Quality of education in the village based government primary schools is found to be an important determinant for private school choice at the primary level. Villages with better quality of education in the government primary schools observe lower enrollment in private primary schools.

4.2 Heterogeneity in Gender Gap

In this section we shall interact the female dummy with some of the relevant variables to understand the nature of gender gap in private schooling. First we interact it with year dummies. *Figure* 3 shows that the gender discrimination has worsened over time. In 1997-98 when both demand and supply of private schooling was relatively much lower,

the point estimate was almost zero, suggesting that there was no differential treatment between boys and girls. In the subsequent years, with the rise in private schooling, the gender gap in private school choice has also widened. In 2007-08, girls were around five percentage points behind boys in private school enrollment. More recently in 2010-11, the discriminatory behaviour towards girls in the school choice decision has been most striking, with the gap being almost 10 percentage points.

Our main specification includes all children belonging to the age group of 6-19 years. It is quite likely that the effect of gender will not be uniform over this whole age band. Therefore, in the next excercise, I allow the interaction between gender and age of the child in the specification. It is found that the gender gap is higher in the early ages. This suggests that households discriminate against girls more while choosing private schools at the primary and middle levels as compared to secondary level. Further, *Figure* 4 shows that this practice of gender discrimination around the early school going age of the child is becoming more prominent over time.

In the next step, we want to see how the gender gap varies with household wealth. If it is the constraint in income that does not allow households to send both boys and girls to fee-charging private schools, and if because of this they strategically select boys for private schooling, then the gender gap should be lower for richer household. On the contrary to this notion, *Figure* 5 shows that the female disadvantage in private school choice deteriorates further as the land holding of the household increases. This implies that it may not be due to income constraint that female children are not sent to private schools, rather the discriminatory preferences of the households come into play in this context.

5 Conclusion

This paper estimates correlated unobserved effects models that also correct for potential sample selection bias using a three-period longitudinal data at the household level to identify the presence of intra-household gender discrimination in private school choice. Our result suggests that there is around six percentage points (or almost 20 percent on average) gender gap in private school enrollment. In the era when overall gender gap in enrollment is reducing, this finding suggests that households choose to provide their sons with what they perceive to be better quality education while ignoring the quality of education received by their daughters. The result also implies that the current growth in private schools in rural areas are unlikely to contribute in reducing the gender gap in education over time. Under the societal system where households value their son's education more than their daughter's, government schools need to focus more on the quality issue to ensure provision of equitable educational opportunity for both boys and girls.

References

[1] Alderman, Harold, Peter F. Orazem and Elizabeth M. Paterno (2001), "School Quality, School Cost and the Public/Private School Choices of Low-Income Households in Pakistan", *Journal of Human Resources*, 36(2): 304-326.

[2] Azam, Mehtabul and Geeta Gandhi Kingdon (2013), "Are Girls the Fairer Sex in India? Revisiting Intra-Household Allocation of Education Expenditure", *World Development*, 42: 143-164.

[3] Chamberlain, Gary (1980), "Analysis of Covariance with Qualitative Data", *Review of Economic Studies*, 47(1): 225-238.

[4] Das Gupta, Monica (1987), "Selective Discrimination against Female Children in Rural Punjab, India", *Population and Development Review*, 13(1): 77-100.

[5] Deolalikar, Anil B. (1993), "Gender Differences in the Returns to Schooling and in School Enrollment Rates in Indonesia", *Journal of Human Resources*, 28(4): 899-932.

[6] Desai, Sonalde, Amaresh Dubey, Reeve Vanneman, Rukmini Banerji (2008), "Private Schooling in India: A New Educational Landscape", in *India Policy Forum*, edited by Suman Bery, Barry Bosworth and Arvind Panagariya, Sage, New Delhi.

[7] Dustmann, C and M. E. Rochina-Barrachina (2007), "Selection Correction in Panel Data Models: An application to the estimation of females' wage equation", *Econometrics Journal*, 10: 263-293.

[8] Garg, Ashish and Jonathan Morduch (1998), "Sibling Rivalry and the Gender Gap: Evidence from Child Health Outcomes in Ghana", *Journal of Population Economics*, 11 (4): 471-493.

[9] Harma, Joanna (2011), "Low cost private schooling in India: Is it pro poor and equitable?", *International Journal of Educational Development*, 31(4): 350-356.

[10] Heckman, James J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1): 153-161.

[11] Kingdon, Geeta (1996), "The quality and efficiency of private and public education: A case-study of urban India", *Oxford Bulletin of Economics and Statistics*, 58(1): 57-82.

[12] Lancaster, Tony (2000), "The incidental parameter problem since 1948", *Journal of Econometrics*, 95(2): 391-413.

[13] Leclercq, Francois (2001), "Household Structures and School Participation: Is There 'Sibling Rivalry' in Rural North India?", *Mimeo.*

[14] Maitra, Pushkar, Sarmistha Pal and Anurag Sharma (2011), "Reforms, Growth and Persistence of Gender Gap: Recent Evidence from Private School Enrolment in India", *IZA Discussion Paper No. 6153.*

[15] Mundlak, Y. (1978), "On the Pooling of Time Series and Cross Section Data", *Econometrica*, 46(1): 69-86.

[16] Muralidharan, Karthik and Michael Kremer (2008), "Public and Private Schools in Rural India", in *School Choice International*, edited by Paul Peterson and Rajashri Chakrabarti, MIT Press, Cambridge, Massachusetts.

[17] Neyman, J. and Elizabeth L. Scott (1948), "Consistent Estimates Based on Partially Consistent Observations", *Econometrica*, 16(1): 1-32.

[18] Ota, Masako and Peter G. Moffatt (2007), "The within-household schooling decision: a study of children in rural Andhra Pradesh", *Journal of Population Economics*, 20(1): 223-239.

[19] Pal, Sarmistha (2010), "Public Infrastructure, Location of Private Schools and Primary School Attainment in an Emerging Economy", *Economics of Education Review*, 29(5): 783–794.

[20] Sawada, Yasuyuki and Michael Lokshin (2009), "Obstacles to school progression in rural Pakistan: An analysis of gender and sibling rivalry using field survey data", *Journal of Development Economics*, 88(2): 335-347.

[21] Tooley, James, Pauline Dixon and S. V. Gomathi (2007), "Private schools and the millennium development goal of universal primary education: a census and comparative survey in Hyderabad, India", *Oxford Review of Education*, 33(5): 539-560.

[22] Woodhead, Martin, Melanie Frost and Zoe 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.

[23] Wooldridge, Jeffrey M. (1995), "Selection correction for panel data models under conditional mean independence assumptions", *Journal of Econometrics*, 68(1): 115-132.

[24] Wooldridge, Jeffrey M. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Massachusetts.

Figure 1: Enrollment rates (%) by gender (6-19 years)

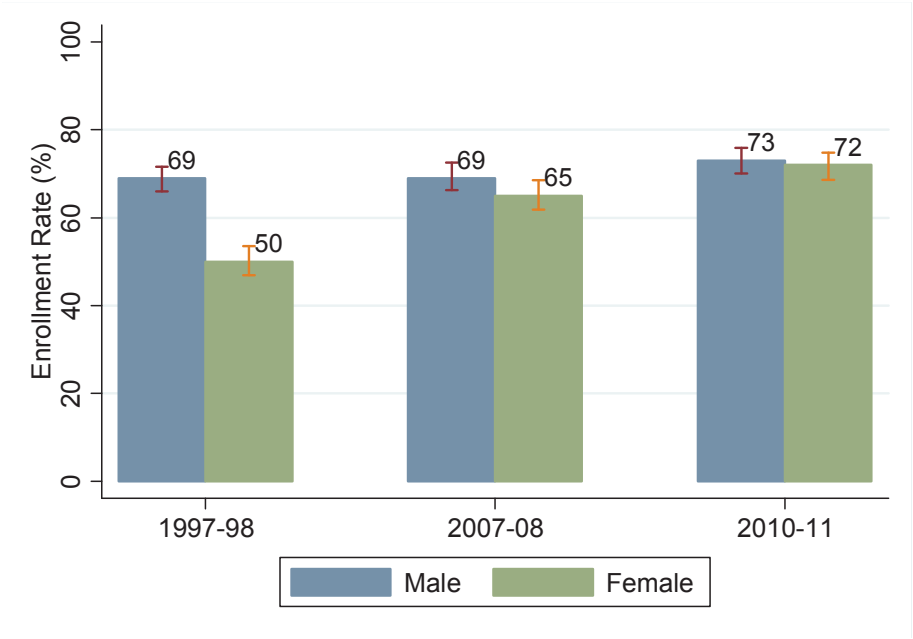


Figure 2: Private school enrollment rates (%) for enrolled children by gender (6-19 years)

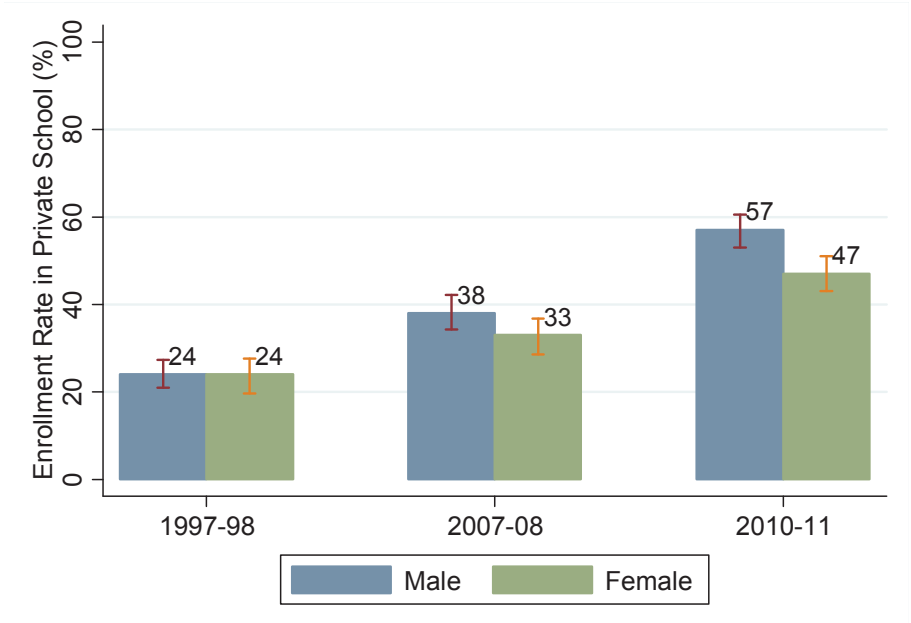


Table 1: Gender gap in private school enrollment – overall results for 6-19 years age group

Variables	Selection corrected	No selection correction
Female	-0.061***	-0.049***
	(0.015)	(0.014)
Age	0.071***	0.012
	(0.020)	(0.013)
Square of age	-0.003***	-0.000
	(0.001)	(0.001)
Birth order	-0.028**	-0.029**
	(0.012)	(0.012)
Mother literate dummy	0.015	0.020
	(0.048)	(0.049)
Father literate dummy	0.001	-0.021
	(0.047)	(0.048)
Head literate dummy	-0.029	-0.033
	(0.061)	(0.063)
Head female dummy	0.039	0.033
	(0.086)	(0.086)
Total number of female children (6-19)	-0.011	-0.009
	(0.015)	(0.015)
Total number of male children (6-19)	-0.002	-0.001
	(0.016)	(0.016)
Household size	0.002	0.004
	(0.007)	(0.007)
Land owned (acre)	0.000	0.000
	(0.003)	(0.004)
Ratio of number of private to total schools	0.139**	0.119*
	(0.065)	(0.065)
Quality index for government primary schools	-0.009	-0.009
	(0.009)	(0.009)
Lambda_1997	0.066	-
	0.065	
Lambda_2007	(0.058)	-
	0.246***	
Lambda_2010	(0.073)	-
	0.225***	
Observations	3,470	3,470
R-squared	0.229	0.225
District-by-time Fixed Effects	Yes	Yes
Correlated Unobserved Effects	Yes	Yes

Bootstrapped (clustered at the household level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Gender gap in private school enrollment: Primary and Post-Primary levels

Variables	Primary (6-10 years)	Post Primary (11-19 years)
Female	-0.091***	-0.043*
	(0.018)	(0.025)
Age	0.125	0.233***
	(0.087)	(0.072)
Square of age	-0.006	-0.008***
	(0.005)	(0.003)
Birth order	-0.011	-0.038*
	(0.019)	(0.021)
Mother literate dummy	-0.006	0.010
	(0.072)	(0.080)
Father literate dummy	0.082	0.003
	(0.056)	(0.059)
Head literate dummy	0.083	-0.036
	(0.074)	(0.093)
Head female dummy	0.053	0.025
	(0.113)	(0.111)
Total number of female children (6-19)	-0.033	-0.014
	(0.030)	(0.029)
Total number of male children (6-19)	-0.011	0.023
	(0.023)	(0.022)
Household size	-0.002	0.004
	(0.010)	(0.010)
Land owned (acre)	0.002	-0.001
	(0.004)	(0.005)
Ratio of number of private to total schools	0.198**	0.061
	(0.094)	(0.089)
Quality index for government primary schools	-0.038***	0.011
	(0.015)	(0.014)
Lambda_1997	0.222***	-0.027
	(0.083)	(0.065)
Lambda_2007	0.056	0.173**
	(0.124)	(0.075)
Lambda_2010	0.329***	0.168*
	(0.119)	(0.095)
Observations	1,573	1,897
R-squared	0.278	0.210
District-by-time Fixed Effects	Yes	Yes
Correlated Unobserved Effects	Yes	Yes

Bootstrapped (clustered at the household level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 3: Estimated gender gap in private school choice, over years

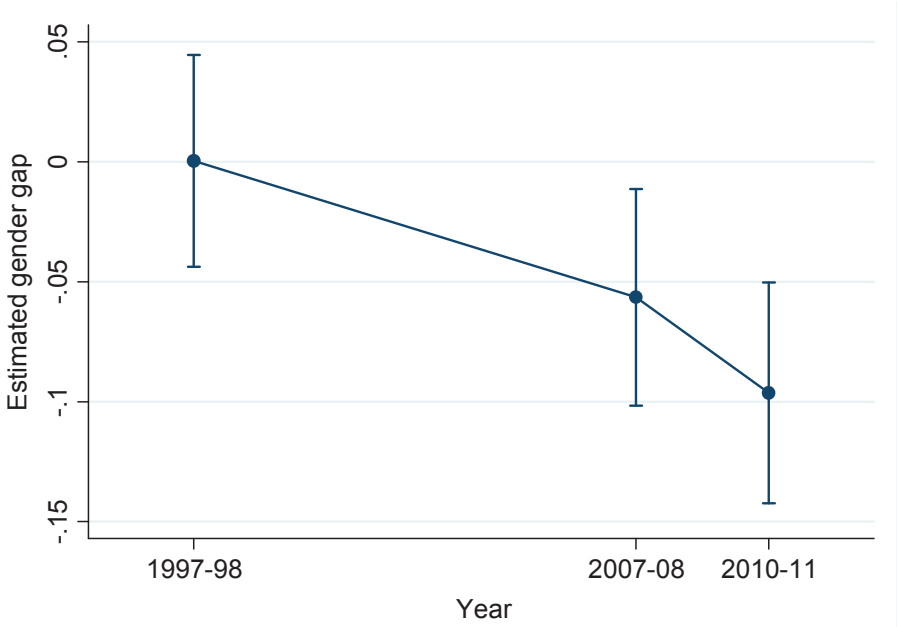


Figure 4: Estimated gender gap in private school choice, over age and year

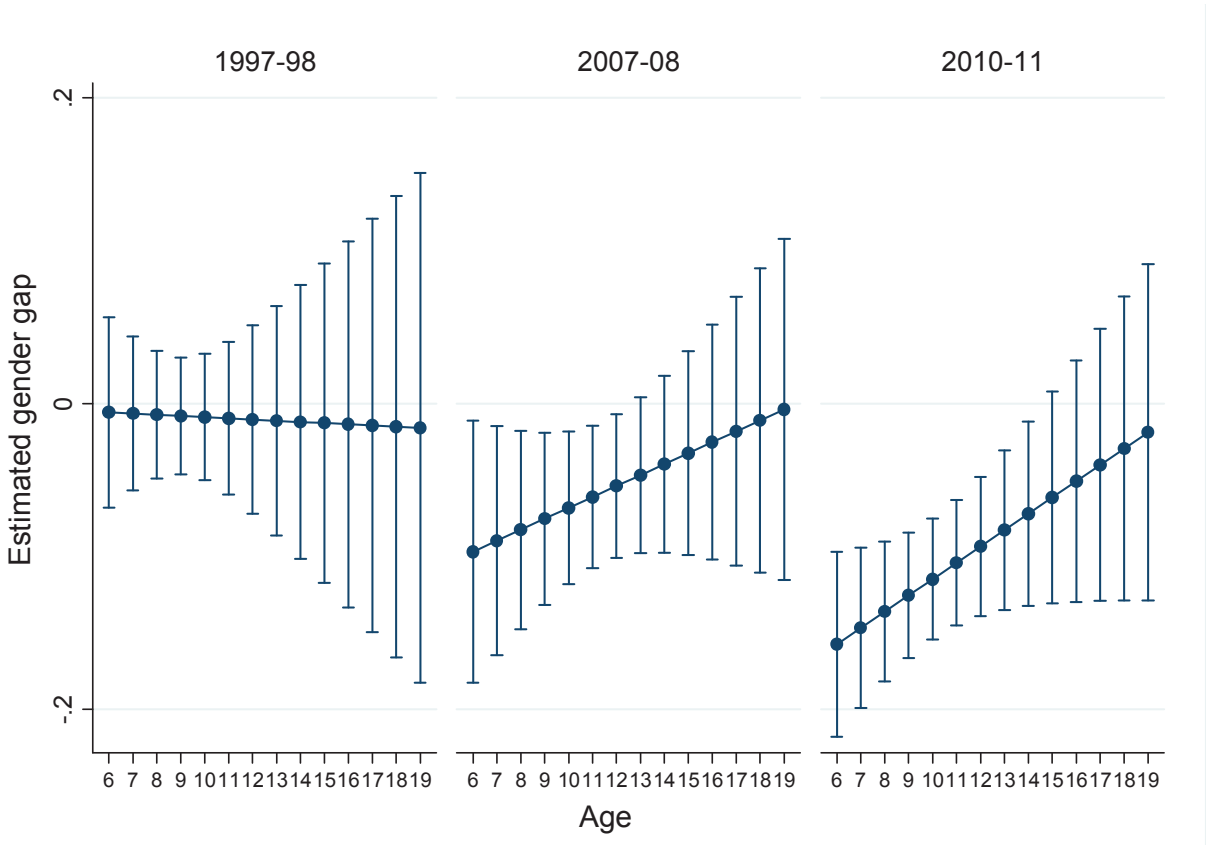
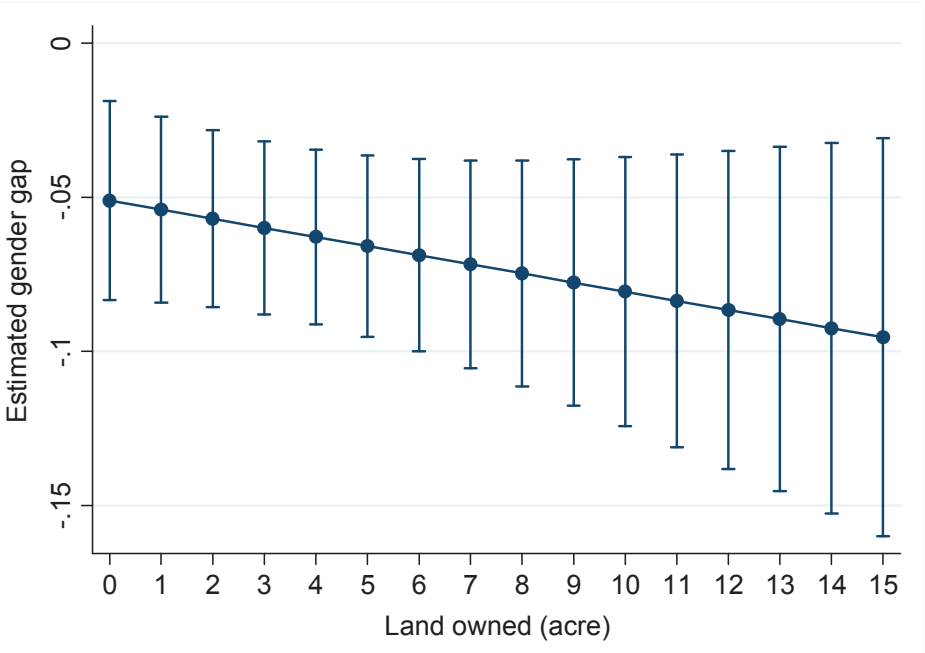


Figure 5: Estimated gender gap in private school choice, over land ownership



Appendix Table A1: Summary statistics

	1997-98			2007-08			2010-11		
Variables	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
Enrollment	1893	0.60	0.49	1557	0.68	0.47	1759	0.72	0.45
Private school enrollment	1893	0.14	0.35	1557	0.24	0.43	1759	0.38	0.49
Private school enrollment | enrollment = 1	1140	0.24	0.43	1051	0.36	0.48	1268	0.52	0.50
Female	1893	0.46	0.50	1557	0.48	0.50	1759	0.48	0.50
Age	1893	11.64	3.86	1557	12.66	3.96	1759	12.48	4.04
Birth order	1893	2.32	1.52	1557	2.25	1.30	1759	2.41	1.57
Mother literate dummy	1893	0.15	0.36	1557	0.19	0.40	1759	0.24	0.43
Father literate dummy	1893	0.50	0.50	1557	0.56	0.50	1759	0.61	0.49
Head literate dummy	1893	0.45	0.50	1557	0.47	0.50	1759	0.52	0.50
Head female dummy	1893	0.04	0.20	1557	0.09	0.29	1759	0.07	0.25
Total number of female children (6-19)	1893	1.72	1.35	1557	1.73	1.17	1759	1.94	1.52
Total number of male children (6-19)	1893	1.99	1.31	1557	1.83	1.13	1759	1.97	1.28
Household size	1893	8.76	4.66	1557	7.93	3.05	1759	9.18	4.47
Land owned (acre)	1893	3.63	7.10	1557	2.41	4.89	1759	3.01	7.25
Ratio of number of private to total schools	1893	0.21	0.35	1557	0.13	0.27	1759	0.22	0.27
Quality index for government primary schools	1893	-1.55	2.42	1557	0.85	1.52	1759	1.05	1.21

Source: SLC data