

IZA DP No. 5706

**Are Girls the Fairer Sex in India? Revisiting
Intra-Household Allocation of Education Expenditure**

Mehtabul Azam
Geeta Kingdon

May 2011

Are Girls the Fairer Sex in India? Revisiting Intra-Household Allocation of Education Expenditure

Mehtabul Azam

*World Bank
and IZA*

Geeta Kingdon

*IoE, University of London
and IZA*

Discussion Paper No. 5706
May 2011

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 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.

ABSTRACT

Are Girls the Fairer Sex in India? Revisiting Intra-Household Allocation of Education Expenditure *

This paper revisits the issue of the intra-household allocation of education expenditure with the recently available India Human Development Survey which refers to 2005 and covers both urban and rural areas. In addition to the traditional Engel method, the paper utilizes a Hurdle model to disentangle the decision to enroll (incur any educational expenditure) and the decision of how much to spend on education, conditional on enrolling. Finally the paper also uses household fixed effects to examine whether any gender bias is a within-household phenomenon. The paper finds that the traditional Engel method often fails to pick up gender bias where it exists not only because of the aggregation of data at the household-level but also because of aggregation of the two decisions in which gender can have opposite signs. It is found that pro-male gender bias exists in the primary school age group for several states but that the incidence of gender bias increases with age – it is greater in the middle school age group (10-14 years) and greater still in the secondary school age group (15-19 years). However, gender discrimination in the secondary school age group 15-19 takes place mainly through the decision to enroll boys and not girls, and not through differential expenditure on girls and boys. The results also suggest that the extent of pro-male gender bias in educational expenditure is substantially greater in rural than in urban areas. Finally, our results suggest that an important mechanism through which households spend less on girls than boys is by sending sons to fee-charging private schools and daughters to the fee-free government-funded schools.

JEL Classification: I21, J16, J71

Keywords: gender bias, educational expenditure, hurdle model, school choice, India

Corresponding author:

Mehtabul Azam
World Bank
1818 H ST, NW
H11-206, Mail Stop: H11-1101
Washington, DC 20433
USA
E-mail: mazam1@worldbank.org

* The views expressed in this paper are entirely of the authors and do not necessarily represent the views of the institutions the authors are affiliated to. All the errors remain of the authors.

1 Introduction

The instrumental case for girls' education is often regarded as too compelling to require argumentation. There are several weighty testimonies in its favour. These comprise the importance ascribed to girls' education in economic growth (Abu-Ghaida and Klasen, 2004; Birdsall, Ross and Sabot, 1993), the significance of female education in improving both market productivity and valued non-market outcomes such as health, nutrition, longevity, fertility and child learning outcomes (Foster and Rosenzweig, 2001; Schultz, 2002; McMahon, 2002; King and Hill, 1993) and, probably reflecting these, the inclusion of gender equality in education as one of the eight Millennium Development Goals. Investing in girls' education has been hailed as possibly the highest return investment available in a developing country (World Bank, 1994). Despite these advantages, however, girls continue to face inferior educational opportunities in many parts of the world, including historically in India.

If girls have inferior educational outcomes vis-à-vis boys this could be due to gender bias in the schooling system, or due to pro-male gender discrimination in the labour market reducing girls' incentive for schooling. An alternative and potentially powerful explanation could also be parental pro-male bias in education manifested in lower intra-household educational expenditure on girls than boys.

Previous work on gender differentials in within-family education expenditure allocation in India (Subramanian and Deaton, 1991, using National Sample Survey data for year 1983) finds evidence of pro-male bias in rural Maharashtra in the 10–14 year age group. Lancaster, Maitra and Ray (2008) use similar data for ten years later and also find significant pro-male bias in the 10-16 year (i.e. upper primary and secondary) age group in rural Bihar and rural Maharashtra. Such biases are not found for urban areas or for primary school children aged 6 to 9. While Subramanian and Deaton concentrate on only one Indian state (Maharashtra), Lancaster, Maitra and Ray's sample is restricted to four Indian states: Bihar, Uttar Pradesh, Kerala, and Maharashtra. In addition, both of these studies use household-level data on the budget share of education in total household expenditure. Kingdon (2005)

uses individual-level education expenditure data from rural India referring to 1993. She finds pro-male bias in the household decision to enroll (or not enroll) children in school, but no evidence of gender bias in education expenditure conditional on enrolling both girls and boys in school. She argues that any gender bias in education within the household can potentially occur in two different decisions : (a) the decision of whether to enroll/retain both sons and daughters in school and, (b) conditional on enrolling both genders, the decision of how much to spend on their schooling. She concludes that a plausible explanation why previous studies did not find intra-household gender bias in education expenditure allocation is because they model the enrolment and conditional expenditure decisions together; since there is pro-male bias in the enrollment decision and no bias (or even a slight pro-female bias) in the conditional education expenditure decision, averaging across these decisions leads to a failure to detect overall gender bias that does exist in the positive education expenditure (enrollment) decision.

However, Kingdon used data from 1993 and since then the liberalization of the Indian economy led to rapid economic growth, reduced poverty and greatly increased school enrolment rates. As such, the gender gap in education outcomes is likely to have fallen. Secondly, Kingdon had household expenditure data only on food, health and education rather than *total* household expenditure. Thus, it worked with the education budget *sub-share* rather than with education budget share, leading to potential problems of bias. Thirdly, Kingdon's work was confined to analysis of bias in rural areas only as her data were exclusively rural. The data used in this paper are from both urban and rural areas. We revisit the issue of the intra-household allocation of expenditure with a new data set collected in 2004-05. We are able to: use better data, test whether there exists bias in urban areas, and ask whether the extent and nature of intra-household gender bias changed between 1993 and 2005 in rural India.

The findings of the paper are as follows. First, there is imperfect correspondence between the results using household-level data and those using individual-level data. We do find

evidence of gender-bias in some states using the household-level data and the traditional Engel method; however, using the same methodology with individual child-level data helps us to unravel gender-bias in many more states. Second, unpacking education expenditure decisions into two parts—a) the decision to enroll in a school, and b) the decision to how much to spend conditional on enrolling—provides additional insights into gender bias since in many states the direction of observed gender bias is opposite in the two decisions. Third, significant progress in gender equality in education has been achieved in rural India between 1993 and 2005: the incidence of gender bias is observed in fewer states in 2005 than was the case in the 1993. We find little evidence of gender-bias in enrollment in age group 5-9, while pro-male bias in conditional education expenditure is observed only in a few states. However, pro-male bias in enrollment is observed in age group 10-14 and 15-19. While in the age group 10-14 pro-male bias in education expenditure is observed in more states than pro-male bias in enrollment, in age group 15-19, pro-male bias is observed in the decision to enroll in more states than the decision on amount of expenditure conditional on enrollment. Fourth, the results reveal a great deal of regional disparities in nature and existence of gender-bias, and between the age groups.

Thus the near achievement of universal enrollment at the elementary level in India has not been translated into higher enrollment at the secondary and senior secondary levels, especially in the rural areas. Not only there exists a large gender gap in enrollment in age group 15-19 in rural areas of many states, the overall enrollment rate in these states remains extremely low, especially for girls. Given that significant pro-male bias is observed in enrollment decision in higher age groups, promoting girls' education remains a priority area.

The paper is organized as follows: Section 2 presents the empirical strategy, Section 3 describes the data, Section 4 presents the results, and Section 5 concludes..

2 Empirical Strategy

We begin the analysis with the estimation of a standard Engel curve linking budget shares on educational expenditure with total household expenditure and the demographic composition of the household. We use the Working-Leser specification as follows:

$$w_i = \alpha + \beta \cdot \ln\left(\frac{x_i}{n_i}\right) + \lambda \cdot \ln(n_i) + \sum_{k=1}^{K-1} \theta_k \cdot \left(\frac{n_{ki}}{n_i}\right) + \psi \cdot z_i + \epsilon_i \quad (1)$$

where w_i is the budget share of education of the i^{th} household; x_i is the total expenditure of the household; n_i is the household size; $\ln\left(\frac{x_i}{n_i}\right)$ is the natural log of total per capita expenditure; $\frac{n_{ki}}{n_i}$ is the fraction of the household members in the k^{th} age-gender class where $k = 1, 2, \dots, K$ refers to the k^{th} age-gender class within household i ; z_i is a vector of other household characteristics such as household head's education, gender, occupation and dummy variables to capture state etc (these variables are defined in the Appendix Table A1); and ϵ_i is the error term. $\alpha, \beta, \lambda, \theta_k$ and ψ are the parameters to be estimated. The Working-Leser specification is relaxed to allow for non-linearity in log per capita expenditure (lnpce). The term n_i allows for an independent scale effect of household size. Since the $\frac{n_{ki}}{n_i}$ fractions add up to unity, one of them has to be omitted from the regression. We allow for 14 age-gender group: males and females aged 0-4, 5-9, 10-14, 15-19, 20-24, 25-60 and 61 and above (omitting the fraction of women aged 61 and above in the regression analysis).¹ The θ_k coefficients capture the effect of household composition on household budgetary allocations. These coefficients tell us the effect of changing household composition holding household size constant, for example by replacing a child aged 5-9 by a child aged 10-14 or by replacing a male with a female in a given age category. The difference across gender can be tested using an F -test under the following null hypothesis:

$$\theta_{kM} = \theta_{kF} \quad (2)$$

¹These age-gender categories are defined as M0to4, F0to4, M5to9, F5to9 etc., and are the proportion of Males (M) and Females (F) in age 0-4, 5-9 and so in a given household.

where M denotes males and F denotes females and k refers to a given age-category.²

Conventionally, Equation 1 has been estimated using OLS including all households. This is because some or much of the bias against girls may occur in the decision of whether to enroll a child in school, i.e. in the zero-versus-positive spending decision, $w_i = 0$ vs $w_i > 0$, rather than only in the decision of how much to spend conditional on enrollment. However, a simple application of the OLS model to data yields parameter estimates which are biased downwards because of censoring of dependent variable as a large proportion of households report zero education expenditure (Deaton, 1997). In addition to biased estimates, averaging across two decisions (enrollment decision and conditional education expenditure decision) leads to a failure to detect overall gender bias if they work in opposite directions. Hence, it is important to separate the two decisions.

We use a Hurdle model (Wooldridge, 2002, p536-38) to separate the initial decision of $w = 0$ from the decision of how much w given positive w .³ Hurdle Models are two-tier models because the hurdle or first tier is the decision of whether to choose a positive w or not ($w = 0$ versus $w > 0$), and the second tier is the decision of how much to spend conditional on spending a positive amount ($w|w > 0$). A simple Hurdle model can be written as follows:

$$P(w = 0|x) = 1 - \Phi(x\gamma) \tag{3}$$

$$\log(w)|x, w > 0 \sim Normal(x\beta, \sigma^2) \tag{4}$$

where w is the share of family budget spent on education, x is a vector of explanatory variables, γ and β are parameters to be estimated while σ is the standard deviation of w .

Equation 3 shows the probability that w is positive or zero, and Equation 4 stipulates that

²For example, testing whether boys aged 10-14 are treated differently from girls aged 10-14, we simply test whether the coefficient on M10to14 (proportion of males aged 10 to 14 years in the household) is significantly different from the coefficient on F10to14 (proportion of females aged 10 to 14 years in the household).

³Tobit model is another available alternative, however, it is identified only if the assumptions of normality and homoskedasticity are fulfilled (Deaton, 1997). Moreover, it assumes that a single mechanism determines the choice between $w_i = 0$ vs $w_i > 0$ and the amount of w given $w > 0$. In particular, $\frac{\partial(P(w>0|x))}{\partial x_i}$ and $\frac{\partial E(w>0|x, w>0)}{\partial x_i}$ are constrained to have the same sign

conditional on $w > 0$, $w|x$ follows a lognormal distribution.⁴ One can obtain an estimate of γ from a probit using $w = 0$ versus $w > 0$ as the binary response. Because of the assumption that conditional on $w > 0$, $\log(w)$ follows a classical linear model, the OLS estimator $\hat{\beta}$ is consistent, and the consistent estimator of σ is just the usual standard error from the OLS regression. The conditional expectation of $E(w|x, w > 0)$ and the unconditional expectation $E(w|x)$ are easy to obtain using properties of log normal distribution:

$$E(w|x, w > 0) = \exp\left(x\beta + \frac{\sigma^2}{2}\right) \quad (5)$$

$$E(w|x) = \Phi(x\gamma) \cdot \exp\left(x\beta + \frac{\sigma^2}{2}\right) \quad (6)$$

which can be easily estimated given $\hat{\beta}, \hat{\sigma}^2, \hat{\gamma}$. One can obtain the marginal effect of x_j on w by transforming the marginal effect of $\log(w)$ and using the exponent. Taking the derivative of the conditional expectation of w with respect to x_j , we can obtain the marginal effect of x on w in the OLS regression of $\log(w)$ conditional on $w > 0$. This is as follows:

$$\frac{\partial E(w|x, w > 0)}{\partial x_j} = \beta_j \cdot \exp\left(x\beta + \frac{\sigma^2}{2}\right) \quad (7)$$

The combined marginal effect of x_j on w , i.e. taking account of the effect of x_j on the probability that $w > 0$ and on the size of $w|w > 0$, can be obtained by taking the derivative of the unconditional expectation of w with respect to x_j . We can use the product rule and take the derivative of the unconditional expectation to obtain the combined marginal effect as follows:

$$\begin{aligned} \frac{\partial E(w|x)}{\partial x_j} &= \gamma_j \cdot \phi(x\gamma) \cdot \exp\left(x\beta + \frac{\sigma^2}{2}\right) + \Phi(x\gamma) \cdot \beta_j \cdot \exp\left(x\beta + \frac{\sigma^2}{2}\right) \\ &= [\gamma_j \cdot \phi(x\gamma) + \Phi(x\gamma) \cdot \beta_j] \cdot \exp\left(x\beta + \frac{\sigma^2}{2}\right) \end{aligned} \quad (8)$$

⁴In our data, the conditional education budget share is indeed log normally distributed.

In the analysis that follows, we estimate three equations for each state: 1) Unconditional OLS equation of the budget share of education (conventional Engel curve) in the household level analysis, and OLS equation of unconditional education expenditure in the individual level analysis; 2) Probit equation of the binary decision whether the budget share of education is positive at the household level analysis, and the probit equation of whether any positive educational expenditure is incurred on the child in the individual level analysis; 3) Conditional OLS of log of budget share of education in the household level analysis, i.e. conditional on positive budget share of education, and OLS of log of conditional education expenditure in the individual level analysis.

Household level equations are fitted for households with at least one child aged 5-19 years. At the individual level we estimate the same equations but, instead of the dependent variable in the OLS equations being the budget share of education (as in household level analysis), the dependent variable is education expenditure on the individual child. Also, all the independent variables are the same in household and individual-level equations except for gender: while household level equations include proportion of household members in 14 age-gender categories, individual level equations use age of child and a dummy variable for male.

3 Data

We use data from the 2005 India Human Development Survey (IHDS-II), a nationally representative household data set collected by the National Council of Applied Economic Research in New Delhi and the University of Maryland (Desai, Reeve and NCAER 2009). IHDS-II covers 41,554 households located throughout India.^{5,6} The data have detailed information

⁵The survey covered all the states and union territories of India except Andaman and Nicobar; and Lakshadweep, two union territories which together account for less than .05 percent of India's population. The data recently became available to public and available from the Data Sharing for Demographic Research program of ICPSR, the Inter-university Consortium for Political and Social Research.

⁶Kingdon (2005) used IHDS-I, which was conducted by the National Council of Applied Economic Research in 1993-94 in rural areas of 16 major states in India. In 2001, the states of Chhattisgarh, Jharkhand,

on education expenditure for persons who are enrolled and comprehensive information on total household expenditure.⁷ The household level analysis is limited to the households with at least one member in age group 5-19, which reduced the sample size to 30,351 households: 19,931 residing in rural area and 10,420 residing in urban areas. The individual level analysis is based at the level of individual child, i.e. on 71,567 children in age group 5-19, out of which 48,882 reside in rural areas and 22,685 reside in urban areas. In our sample, the share of education expenditure in household expenditure is 3.5 percent in rural areas, 6.4 percent in urban areas, and 4.3 percent in all India sample.⁸

Table 1 presents the state wise current enrollment rate for boys and girls. Among the rural age group 5-9, girls have statistically significantly lower enrollment than boys in only two states (Rajasthan and Madhya Pradesh) which is a substantial improvement from the 1993 situation when this was the case in nine of the 16 major states examined (Kingdon, 2005). In urban areas, only Jharkhand has lower enrollment for girls in age group 5-9, while in most of the other states girls have comparable or better enrollment than boys. Between 1993 and 2005, there is also a significant improvement in enrollment for girls in the rural age group 10-14: girls have lower enrollment than boys in only seven states in 2005 compared to the 1993 situation where girls' lower enrollment was observed in all the states except Kerala and West Bengal. A significant improvement over time in gender parity in enrollment is also observed in age group 15-19. While rural girls had lower enrollment than boys in almost all major states (except Kerala) in 1993, by 2005 they had lower enrollment only in half of the states. In urban areas, only four states have lower enrollment for girls in age groups 10-14 and 15-19; however it is surprising that lower enrollment of girls than boys is observed in Gujarat (in both these age groups) and in Maharashtra (in the 10-14 age group) which are

and Uttarakhand were carved out of the parental states of Madhya Pradesh, Bihar, and Uttar Pradesh. In addition we also have Jammu & Kashmir, North East (Combination of smaller states in North East of India), and Delhi (primarily urban).

⁷Total education expenditure on a school going child is derived by adding expenditure on school fees, books, uniforms, and private tuition.

⁸Note that our sample consists households who have child in age group 5-19. For all households, the share of education expenditure in household expenditure is 2.7 percent in rural areas, 4.9 percent in urban areas, and 3.4 percent in all India.

economically well-to-do states in India.

Table 2 presents mean educational expenditure on all children in the three concerned age groups (it includes zero education expenditure by the non-enrolled), while Table 3 presents conditional education expenditure, i.e. education expenditure for enrolled children only. Table 2 shows that education expenditure on girls is lower compared to boys in a larger number of states in rural areas than in urban areas in all the three age groups. It is interesting that the pro-male bias in education expenditure is observed in far fewer states when we take education expenditure conditional on enrollment. This suggests that some of the gender bias in education occurs at the stage of enrolment itself; a significant part of the gender gap in educational expenditure comes through girls' lower probability of enrollment (since non-enrollment implies zero education expenditure), especially in the higher age groups.

4 Results

The results are divided into two parts. In the first part, we present the results obtained by using household-level aggregated data. We explore, using conventional Engel curve approach, whether there is evidence that the allocation of household education expenditure favors male over female children. We also explore whether averaging across two decisions (positive education expenditure decision and actual education expenditure amount decision) makes a difference to detecting gender bias. As explained in the methodology section, this is done by comparing results from conventional Engel curve approach with the results from the hurdle model. In the second part, we explore whether aggregation of data at the household level hinders detection of gender bias; we do this by comparing results from household-level data with those obtained using individual child level data. With child level data we also ask again whether separating the two decisions (positive education expenditure decision and the actual amount of expenditure decision) enhances our understanding of the nature of gender bias in education. To achieve this, we estimate the unconditional OLS of education

expenditure and also the hurdle model using individual child level data. We further check the robustness of the results obtained from individual-level analysis by introducing household fixed effects. Finally, we explore whether any observed bias in education expenditure comes through differential choice of school-type (private or public) for sons and daughters.

4.1 Analysis with household level data

Using household-level data, we estimate three equations for rural and urban areas separately (and pooled together) for each state, and for all India: (a) a binary probit of whether the household’s education budget share is positive or zero; (b) the OLS of the natural log of education budget share, conditional on positive education budget share; and (c) the conventional Engel curve equation. We also put the results of (a) and (b) together to come up with the combined marginal effect of the gender variable, in the way described above in Equation 8. For space reasons, we do not report the full estimation results (that would entail reporting nearly 200 equations - 3 equations each for rural, urban and pooled (rural plus urban) for each of 22 states), but the results are available from the authors. For brevity, we only report in Table 4a the main result, i.e. the difference in marginal effect (DME) of the demographic variables (proportion of males and females in given age groups within the household), for each of the three age groups: 5-9, 10-14, and 15-19. The DME for a given age group within each state is calculated from the results of the full model for the state. For example, to calculate DME for all-India, we estimate the three equations using the all-India sample. In the probit equation of ‘positive educational expenditure’ (which is a proxy for enrolling children in school, since school enrollment is associated with at least some expenditure), at the all-India level (see Appendix Table A1), the marginal effect of the variable ‘proportion of household males aged 5 to 9’ or M5to9 is 0.639 and the marginal effect of the variable ‘proportion of household females aged 5-9’ or F5to9 is 0.577. Thus the gender difference in the age group 5-9 is 0.062. Table 4a (first cell) shows this difference multiplied by 100, which is 6.18. The F-test (Equation 2) tells us that this

difference in marginal effect (DME) is statistically insignificant at 5% significance level. That is, in the 5-9 age group in India as a whole, there is no statistically significant difference in families' propensity to incur positive educational expenditure for girls and boys. Indeed, this remains the case for all states except Rajasthan where there is a significant difference (at the 10% level) in households' propensity to incur positive educational expenditure on girls and boys: in Rajasthan, when an extra boy is added to the household in the 5-9 age range, the household's probability of incurring positive education expenditure (enrolling children in school) is 24 percentage points higher than when an extra girl is added to the household in that age range.

In the 10-14 age group (the middle school age group), there is clear pro-male bias in this propensity to enroll children in school at the All-India level and in five states - Bihar, Jharkhand, Madhya Pradesh, Rajasthan and Jammu Kashmir. In the 15-19 age group (the secondary school age group), there is again clear pro-male bias in school enrollment probability in All-India and in eight of the states, including the apparently more 'progressive' states such as Karnataka, West Bengal and Maharashtra. Comparing the probit equations across the age-groups then, we see a pattern whereby gender bias in school enrollment (proxied by positive education expenditure) increases with age, the pro-male bias being smallest at the primary school age, larger at the middle school age and largest at the secondary school age.

Looking at education expenditure conditional on household incurring positive education expenditure (columns 2 in Table 4a), we again see a pattern where gender bias increases by age group. Statistically significant pro-male bias in actual education expenditure is found in only two states in the 5-9 age group (Chhattisgarh and Madhya Pradesh), in four states in the 10-14 age group, and in eight states in the 15-19 age group.

Column (3) of Table 4a reports the results of the hurdle model, i.e. it reports the combined marginal effect of the demographic variables (M5to9, F5to9, etc.), putting together the results from the probit equation (column 1) and the conditional OLS equation (column 2). Here again, unsurprisingly, we see a pattern of increasing bias by age group. Moreover,

we see that—taking India as a whole—the combined marginal effect of the gender variable is progressively larger as age group increases: it goes from 1.07 in the 5-9 group, to 1.24 in the 10-14 group and increases to 2.92 in the 15-19 age group.

Finally, Column 4 of Table 4a reports the findings of conventional Engel Curve analysis and this allows us to compare the results with the hurdle model of Column 3. According to Column 4, in the 5-9 age group, there is statistically significant pro-male gender bias in education expenditure (at the 5% level) in only two states whereas the hurdle model (Column 3) shows such bias in three states (and at ‘All-India’). Again, in the 10-14 age group, the conventional analysis shows significant bias in four states (including ‘All-India’) whereas hurdle model shows bias in five states. Finally in the 15-19 age group, conventional analysis shows bias at the 5% level in six states (including ‘All-India’) whereas hurdle model shows it in 9 states. Thus, it is noticeable from Table 4a that the hurdle model is better able to detect gender bias than the conventional Engel curve model, as also found in Kingdon (2005). This is because the hurdle model is a more flexible formulation; it allows for the possibility that gender bias in the enrollment (positive expenditure) decision can differ from any gender bias in the actual education expenditure amount decision.

Unpacking the decisions using hurdle model, we find that in some cases the bias works in opposite directions. For example, in case of Gujarat, a pro-male bias is observed (although not significant) in the enrollment decision (probit, Column 1), however, a pro-female bias is observed in conditional education expenditure (conditional OLS, Column 2). Unpacking the decisions using the hurdle model helps us to unravel gender bias in more states than the traditional Engel method because averaging across the two different decisions mutes the gender bias in Engel method. Hence, the incorrect functional form of Engel Method can be partially blamed for its failure to detect gender bias in some cases; however, the aggregation of data at the household level can also make the detection of gender-bias difficult. In the next section, we explore whether using individual-level data makes any difference as far as detection of gender bias is concerned.

Dividing the analysis by rural (Table 4b) and urban (Table 4c) indicates that the substantial gender biases we observe in Table 4a are driven largely by rural areas. There is little gender bias in urban parts of India, barring a few instances. Moreover, while pro-male bias is observed in the enrollment decision in rural areas of quite a few states, the 2005 situation is a significant improvement over the 1993 situation when pro-male bias is observed in rural areas of 11 out of 16 states (Kingdon, 2005).

4.2 Analysis with individual-level data

Does individual-level expenditure provide a more reliable way of detecting gender bias than using household level data? Since we have educational expenditure information at the level of the individual child (as well as, by aggregation, at the level of the household), it is possible to compare household-level Engel curve results of Section 4.1 with individual-level analysis. In this child level analysis, the dependent variable is education expenditure on the individual child (rather than household budget share of education used in Section 4.1). Moreover, instead of demographic variables such as household's proportion of males aged 5-9 and household's proportion of females aged 5-9, and so forth, the variable of interest is the gender variable (a dummy variable: MALE).⁹ The remaining explanatory variables in the individual level equations are identical to those in the household level equations of Table 4. At the individual child level, we estimated 621 separate equations: (22 states + all India= 23) \times (rural, urban, all= 3) \times (3 age groups) \times (3 equations). For brevity, we only report the marginal effect on the gender variable MALE from these equations in Tables 5a - 5c as our main interest lies in detecting gender bias.

The marginal effects on MALE in Tables 5a-5c are not comparable with the difference in marginal effects (DME) of the household demographic variables reported in Tables 4a-4c. This is because the household demographic variables in a household-level regression are

⁹Since MALE is a discrete variable, the marginal effect of MALE in the combined hurdle model (probit + OLS) is estimated by calculating the expected values of unconditional expenditure in Equation 6 with MALE=1 and with MALE=0, and then taking the difference, rather than by taking derivatives, as in Equation 8.

not identical to the dummy variable MALE in the individual-level regression. In addition, the dependent variable in the conditional and unconditional OLS equations in Table 5 is education expenditure on the individual child but in Table 4 the corresponding dependent variable is household education budget share.

Two things stand out through simple observation of Table 5a (rural plus urban pooled). First, even the unconditional OLS picks up gender bias in many more states than what it picked up at the household level (Table 4a). For example, in the age group 10-14, Table 4a (household level results) shows significant pro-male bias in only four states while Table 5a (individual-level results) shows such bias in thirteen states (including ‘All-India’). This suggests that there is something in the aggregation that makes it more difficult to pick up gender differences in education expenditure. Second, as before with household level data, so also with individual-level data, comparison of hurdle model results (combined probit and conditional OLS) with unconditional OLS results shows that the hurdle model is more effective at picking up gender bias than the conventional unconditional OLS model. For example, in the 5-9 age group in Table 5a, the unconditional OLS results of Column 4 show significant pro-male bias in only 6 states (at the 5% significance level), while the Hurdle model results of Column 3 show such bias in 8 states.

As with household level data, the incidence of pro-male bias in enrollment decision increases in higher age groups and is highest in the age group 15-19 (Table 5a). It is noteworthy that while pro-male bias is observed in conditional education expenditure in many states in age group 10-14, the incidence of such bias in conditional education expenditure is less in age group 15-19. In the age groups 5-9 and 10-14, much of the gender-differentiated treatment occurs at the second stage i.e., in the decision of how much to spend (given that children of both genders are enrolled in school); however, pro-male bias in the enrollment decision also remains important in age group 10-14 years. In contrast, much of the gender-differentiated treatment in age group 15-19 occurs at the stage of the decision whether to even incur positive education expenditure (enroll a child in school)—pro-male bias is observed at the second

stage in only 6 states in comparison to the 13 states where pro-male bias is observed at the first stage. These patterns hold both in urban and rural areas, and are starker in the urban areas (Table 5c).

In some instances, the marginal effect of MALE in the conditional expenditure equation is negative, that is, girls have somewhat higher education expenditure, conditional on being in school, though this pro-female bias is statistically significant only in Tamil Nadu (Table 5a). What is of more concern is that a pro-male bias in enrollment is observed in age group 15-19 in well to do states like Gujarat, Maharashtra, and Karnataka. This raises the question why parents are more reluctant to send a girl child to school at the higher age. One of the reasons for this gender bias could originate on the supply of the market for education. There may be lack of single sex secondary or higher secondary schools for girls in rural areas. People in rural areas might be more reluctant to send an adolescent girl to a coeducational school or to a school which involve significant commuting time. There are also reasons for different demand for education. Providing education at the secondary and senior secondary levels (age group 15-19) involves a significant expenditure in rural areas (Table 3). In this case a economically weak household may prefer to educate sons who are more likely to support them in old age. In addition, child marriages still continue despite the fact that the Child Marriage Restraint Act was enacted as far way back as in 1929. Rajasthan, Madhya Pradesh, Andhra Pradesh, Bihar and Uttar Pradesh have high incidence of child marriages, and girl child are more prone to an early marriage.

4.2.1 Household Fixed Effects: Gender differences within or across households?

Using individual-level data, we found that pro-male bias exist in many states (Section 4.2). However, Jensen (2002) suggests that gender inequality in outcomes could arise even in the absence of any parental bias against daughters. If parents have a strong preference for male children, they will continue child bearing until one (or their desired number of) male offspring is born. This type of fertility behavior implies that, on average, female children will have

a larger number of siblings and larger household size than male children. So any observed lower educational expenditures on girls than boys could be an across-household phenomenon due to differential household sizes for girls and boys in the population. If household size is endogenously chosen in the way Jensen (2002) describes then simply controlling for household size will not be suffice. To check the robustness of our findings, we introduce household fixed effects in the individual-level analysis. Introducing household fixed effects is a powerful way of controlling for unobserved parental fertility preferences and thus for the endogeneity of household size.

We estimate three equations using individual-level data for each state: 1) a Linear Probability Model (LPM) equation of ANYEDEXP (whether any positive expenditure was incurred on the child’s education); 2) an OLS equation of the educational expenditure (EDUEXP) conditional on positive educational expenditure; and 3) a OLS unconditional educational expenditure (EDUEXP) equation. These equations are fitted on the sample of only those households that have at least one child of each gender in the relevant age range.¹⁰ Table 6 reports coefficient of the MALE dummy in the three equations from the household fixed effects estimation. We find pro-male bias in many states in both decisions—the decision to enroll as well as the conditional and unconditional expenditure decisions—in all the three age groups, though there are some variations.

Family fixed effects (FE) results (Table 6) are similar but not identical to the OLS results (Table 5a). For example, the coefficient on MALE gender dummy variable in the unconditional education expenditure equation for age 5-9 group is positive and statistically significant at the 5% level in Andhra Pradesh and Chhattisgarh in Table 5a but insignificant in Table 6. Similarly in the other two age groups there are some differences in results between the OLS and FE approaches of Tables 5a and 6. However, for most states, findings in the OLS individual-level analysis of Table 5a are similar to those in the family FE analysis and thus we take it that most gender differences cannot be simply attributed to differences in

¹⁰The estimation is done using combined sample (urban and rural) in each state because of sample size considerations.

the household size across the population.

4.2.2 Does the bias in educational expenditure through school choice?

In individual-level analysis, we found that statistically significant pro-male bias in education expenditure exists in many states. Is less spent on enrolled girls than boys through differential school-type choice for the two sexes; for example, through a greater likelihood of sending boys to private schools than girls? There are three types of schools in India: government schools, private schools and government-aided schools. In government-aided schools, teachers receive their salary directly from the state and are recruited by a government appointed commission but their routine operations are governed by the private management (Kingdon, 2008). Since in cost and teacher qualification they are similar to government schools, we combined these with government schools.¹¹ The education expenditure is considerably higher for children attending private schools (Appendix Table A2) in all the states. Raw gender differences in private school attendance show that boys are significantly more likely to attend private schools in 5 states in age group 5-9, in 7 states in age group 10-14, and three states in age group 15-19. Only in Orissa, girls in age group 15-19 are more likely to attend private school. At the all India level, while boys are more likely to attend private school in age group 5-9 and 10-14, there is no difference in private school attendance in age group 15-19. At the senior secondary level, the availability of private schools is limited in comparison to availability of private schools at lower levels. This probably is reflected in no difference in private school attendance by gender in age group 15-19.

However, school choice is determined by a number of observed and unobserved factors. To control for observed factors, we estimate a linear probability model (LPM) of school choice on all currently enrolled children, conditioning on observed household characteristics.¹² The gender dummy variable (Male) captures the impact of gender on the choice of private

¹¹About 5.4 percent of school going children reported to be attending government-aided schools.

¹²Independent variables include log of per capita expenditure (lpce), square of lpce, log of household size, education and occupational status of household's head, area of residence, a gender (MALE) dummy, dummies for states (in case of all India).

schooling. We only report the coefficient of gender dummy in Table 7. The estimates from LPM model show that even after controlling for household level observed factors, boys are more likely to attend private schools. The pro-male bias in attendance of private school in age group 15-19 is observed in more states after controlling for observed factors. For all India, no difference in private school attendance is observed in age group 15-19; however, after controlling for observed factors, the boys are 2 percent more likely to attend private school.

To control for both observed and unobserved factors at the household-level, we re-estimate the LPM models with household fixed effects. By controlling for observables and un-observables within the family, fixed-effects estimation allows us to test whether the observed pro-male bias is an across household or within household phenomenon. For this, the sample of households is restricted to only those households who had at least one child of each gender in the relevant age group and currently enrolled in school. In age groups 5-9 and 10-14 boys are five percentage points more likely to be enrolled in private schools than girls. When we allow for household fixed effects, pro-male bias is observed in more states in these age groups. Fixed-effects estimation strengthens the findings that boys are more likely to attend private schools than the girls. Thus one mechanism through which households achieve lower expenditures on education for girls is through a lower probability of sending them to private schools.

5 Conclusion

The paper revisits the issue of gender-bias in educational expenditure in India. Most of the existing literature on the gender bias in education expenditure in India uses data from the early 1990s, and concentrates on only a few states. Kingdon (2005) was the first study to use all India rural data from 1993 to study the gender-bias in education expenditure. However, since then the liberalization of the Indian economy led to rapid economic growth,

reduced poverty and greatly increased school enrolment rates. This paper revisits the issue of gender-bias in educational expenditure using a recently available India Human Development Survey (IHDS-II), which was collected during 2004-05 and is representative at the state and national level. In addition to providing evidence on gender bias in both urban and rural areas of each state, the paper also establishes the progress made in gender-parity in rural India between 1993 and 2005 by comparing the results of this paper with Kingdon (2005), which concentrated only on rural areas because of non-availability of data from urban areas.

We find that Unconditional OLS—which is what the past literature uses—is weak in detecting gender bias. Unpacking the decisions into two parts (through hurdle model) does a much better job. In addition, availability of individual level data greatly improves the ability to detect gender bias—disaggregation of data at the household level appears to mute gender effects and make it harder to detect gender bias. While we find evidence of gender bias in a few states using the household-level data and traditional Engel method, using the same methodology with individual-level data helped us to unravel gender-bias in many more states. We find little evidence of gender-bias in the enrollment decision in age group 5-9 (whether in household-level or in individual-level data), while pro-male bias in education expenditure conditional on enrollment is observed in several states.

The incidence of pro-male bias in enrollment is substantially higher in older age groups. While in the 10-14 age group, pro-male bias in education expenditure is observed in more states than the pro-male bias in enrollment, in age group 15-19, pro-male bias is observed in more states in the enrolment decision than in the amount of expenditure incurred conditional on enrollment. Overall, the results reveal regional disparities in nature and existence of gender-bias, and between the age groups.

Evidence of greater gender-bias in higher age groups raises some important policy issues. Given the large economic returns to higher education in India (Kingdon, 2009; Colclough, Kingdon and Patrinos, 2010), and given the benefits of higher education for economic growth and development, increasing access to higher education is increasingly an important objec-

tive of policy-makers everywhere. Given that significant pro-male bias is observed in the enrollment decision in higher age groups in many economically backward states, promoting girls' education remain an area of priority. What is more of the concern is the existence of pro-male bias in enrollment in age group 15-19 in economically well to do states such as Gujarat and Maharashtra. It remains an important question why parents are reluctant to send girl children to school in higher age groups, specifically in age group 15-19. Whether it is a supply side constraint (non availability of single sex school in rural areas that may be preferred for adolescent girls) or demand side constraints related to lingering cultural and gender norms needs to be further explored.

References

- [1] Abu-Ghaida, D. and Klasen, S. (2004), “The Costs of Missing the Millennium Development Goal on Gender Equity,” *World Development*, 32(7), 1075-1107.
- [2] Ahmad, A. and Morduch, J. (2002), “Identifying Sex Bias in the Allocation of Household Resources: Evidence from Linked Household Surveys from Bangladesh,” *Mimeo*, Department of Economics, New York University.
- [3] Aslam, M. (2009), “The Relative Effectiveness of Government and Private Schools in Pakistan: Are Girls Worse off?” *Education Economics*, 17(3), 329-354.
- [4] Aslam, M. and Kingdon, G. G. (2009), “Gender and Household Education Expenditure in Pakistan,” *Applied Economics*, 40 (20), 2573-2591.
- [5] Birdsall, N., Ross, D., and Sabot, R. (1993), “Underinvestment in Education: How Much Growth has Pakistan Foregone?,” *The Pakistan Development Review*, 32(4), 453-499.
- [6] Colclough, C., Kingdon, G., and Patrinos, H. (2010), “The Changing Pattern of Wage Returns to Education and its Implications,” *Development Policy Review*, 28 (6), 733-747.
- [7] Deaton, A. and Paxton, C. (1998), “Economies of Scale, Household Size, and the Demand for Food,” *Journal of Political Economy*, 106, 897–910.
- [8] Deaton, A. (1997), “The Analysis of Household Surveys: A Microeconometric Approach to Development Policy,” The Johns Hopkins University Press, Baltimore, Maryland, USA.
- [9] Deaton, A. (1989), “Looking for Boy-Girl Discrimination in Household Expenditure Data,” *World Bank Economic Review*, 3, 1–15.

- [10] Desai, S., Vanneman, R., and National Council of Applied Economic Research. (2005), “India Human Development Survey (IHDS),” Computer file, ICPSR22626-v5. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2009-622.doi: 10.3886/ICPSR22626.
- [11] Drèze, J. and Kingdon, G. G. (2001), “Schooling Participation in Rural India,” *Review of Development Economics*, 5 (1), 1-24.
- [12] Foster, A. and Rosenzweig, M. (2001) “Missing Women, the Marriage Market, and Economic Growth”, *Mimeo*, Brown University.
- [13] Haddad, L. and Reardon, T. (1993), “Gender Bias in the Allocation of Resources Within Households in Burkina Faso: A Disaggregated Outlay Equivalent Analysis,” *The Journal of Development Studies*, 29, 260–276.
- [14] Himaz, R. (2010), “Intrahousehold Allocation of Education Expenditure: The Case of Sri Lanka, *Economic Development and Cultural Change*,” 58 (2), 231-258.
- [15] Jensen, R. (2002), “Equal Treatment, Unequal Outcomes? Generating Sex Inequality Through Fertility Behaviour,” *Mimeo*, Harvard University, John F. Kennedy School of Government.
- [16] King, E. and Hill, M. (1993), “Women’s Education in Developing Countries: Barriers, Benefits and Policies,” The John Hopkins University Press, Washington, D.C.
- [17] Kingdon, G. (2009), “The Economic Benefits of Education in India”’, *Journal of the Lucknow Management Association*, 4th Annual Issue, November.
- [18] Kingdon G. G. (2008), “Private and Public Schooling: The Indian Experience,” in Chakrabarti, R. and Petersen, P. edited: *School Choice International: Exploiting Public-Private Partnerships*, MIT Press.

- [19] Kingdon, G. G. (2007), “The Progress of School Education in India,” *Oxford Review of Economic Policy*, 23 (2), 168–195.
- [20] Kingdon, G. G. (2005), “Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure,” *Economic Development and Cultural Change*, 53, 409–451.
- [21] Kingdon, G. G. (1998), “Does the Labour Market Explain Lower Female Schooling in India?” *Journal of Development Studies* 35, 39–65.
- [22] Lancaster, G., Maitra, P., and Ray, R. (2008), “Household Expenditure Patterns and Gender Bias: Evidence from Selected Indian States,” *Oxford Development Studies*, 36(2), 133-157.
- [23] McMohan, W.W. (2002), “Education and Development: Measuring the Social Benefits,” Oxford University Press.
- [24] Schultz, P. T. (2002), “Why Governments Should Invest More to Educate Girls,” *World Development*, 30(2), 207-225.
- [25] Subramanian, S. (1995), “Gender Effects in Intra-Household Allocation in India, *Unpublished manuscript*, Department of Economics, Cornell University.
- [26] Subramanian, S. and Deaton, A. (1991), “Gender Effects in Indian Consumption Patterns,” *Sarvekshana*, 14, 1-12.
- [27] Wooldridge, J. M. (2002), “Econometric Analysis of Cross Section and Panel Data,” The MIT Press, Cambridge, Massachusetts, USA.
- [28] World Bank. (1994), “Investing in All the People: Education Women in Developing Countries,” Washington, D.C.

Table 1: Current enrolment rate of children by age group, location and gender

State	Rural									Urban								
	Age 5-9			Age 10-14			Age 15-19			Age 5-9			Age 10-14			Age 15-19		
	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap
All India	76.7	73.8	2.8**	86.7	80.4	6.3**	48.6	34.3	14.2**	85.4	85.7	-0.2	89.8	87.4	2.5**	59.6	56.3	3.3**
Andhra Pradesh	90.3	90.1	0.2	88.2	81.9	6.3**	53.4	30.8	22.6**	94.6	94.2	0.4	90.8	89.1	1.6	60.3	52.3	8.0
Assam	70.3	74.7	-4.4	72.0	78.7	-6.7	35.7	42.6	-6.9	86.8	85.2	1.5	88.2	86.1	2.2	72.6	70.7	1.9
Bihar	59.3	55.0	4.3	80.8	62.5	18.3**	40.0	26.7	13.3**	78.9	80.8	-1.9	85.7	80.6	5.1	57.3	54.7	2.6
Chhattisgarh	66.4	62.6	3.8	88.6	83.2	5.4	46.1	22.7	23.5**	85.3	90.0	-4.7	90.3	83.3	7.0	70.8	62.5	8.3
Delhi	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	82.3	82.1	0.2	92.3	92.1	0.1	62.9	59.9	3.0
Gujarat	84.8	77.3	7.4*	89.7	81.8	7.9**	38.8	17.6	21.2**	83.3	83.3	0.0	89.9	82.9	7.0**	55.1	43.4	11.7**
Haryana	82.3	78.6	3.7	89.9	87.3	2.6	58.1	43.4	14.7**	88.4	90.7	-2.4	85.7	89.3	-3.6	83.6	66.7	16.9**
Himachal Pradesh	88.5	85.6	3.0	96.8	96.0	0.7	77.1	74.8	2.3	94.7	98.1	-3.4	95.5	97.2	-1.7	84.1	83.3	0.7
Jammu & Kashmir	93.4	91.1	2.3	92.1	88.4	3.7	62.5	55.7	6.8	94.6	92.9	1.7	96.5	93.6	2.9	85.6	72.7	12.8**
Jharkhand	61.2	66.7	-5.5	78.9	64.6	14.3**	32.1	35.3	-3.2	91.0	77.9	13.1**	91.0	89.2	1.9	76.3	64.2	12.1**
Karnataka	75.4	72.8	2.6	87.3	87.2	0.1	49.7	37.5	12.1**	88.2	87.8	0.4	92.1	92.7	-0.6	59.2	57.0	2.2
Kerala	81.1	94.1	-13.0**	99.6	100.0	-0.4	75.2	73.9	1.3	79.8	79.4	0.4	99.2	99.2	0.0	77.1	69.1	8.0
Madhya Pradesh	71.4	65.5	6.0**	87.4	79.7	7.7**	51.7	28.8	22.8**	82.8	82.9	-0.1	91.4	83.6	7.8**	54.3	51.6	2.8
Maharashtra	77.5	78.4	-0.9	92.7	87.6	5.1*	54.1	41.7	12.4**	83.4	83.3	0.1	97.1	93.7	3.3**	60.4	60.8	-0.4
Northeast	82.2	85.7	-3.6	93.1	86.4	6.6*	58.0	52.4	5.6	90.5	99.1	-8.6**	94.8	98.0	-3.2	82.3	84.6	-2.3
Orissa	85.7	84.9	0.8	81.8	77.2	4.6	28.7	18.1	10.6**	89.6	85.9	3.7	86.5	87.7	-1.2	47.8	45.9	1.9
Punjab	85.6	87.5	-1.8	91.1	87.1	3.9	52.2	48.5	3.7	93.1	95.7	-2.6	92.6	92.9	-0.3	62.8	71.6	-8.8*
Rajasthan	79.2	65.1	14.0**	86.9	67.8	19.1**	47.4	23.7	23.7**	82.7	80.3	2.4	85.7	79.3	6.3**	55.0	46.6	8.3*
Tamil Nadu	86.0	83.9	2.1	96.1	92.2	3.8	62.8	44.4	18.3	89.8	95.7	-5.9**	94.6	93.9	0.7	61.2	54.0	7.1
Uttar Pradesh	77.9	74.9	3.1	86.7	82.7	4.0**	49.9	31.6	18.3**	81.2	80.4	0.9	78.4	76.4	2.0	46.9	50.0	-3.1
Uttarakhand	79.3	81.1	-1.7	94.2	90.5	3.7	72.7	58.5	14.2*	64.4	89.3	-24.9**	88.3	87.2	1.1	65.9	66.1	-0.2
West Bengal	81.1	77.2	4.0	78.2	79.2	-0.9	37.4	30.6	6.8*	90.5	85.5	5.0	86.1	81.7	4.4	57.9	55.9	2.0

Note: [1]. Gender gap is the difference between male and female enrolment in each age group. Positive gap implies pro-male bias, while negative gap implies pro-female bias. [2]. ** statistically significant at 5% level; *statistically significant at 10% level; N/A: not applicable.

Table 2: Education Expenditure on all children (Enrolled and Non-enrolled) by age group, location and gender

State	Rural									Urban								
	Age 5-9			Age 10-14			Age 15-19			Age 5-9			Age 10-14			Age 15-19		
	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap
All India	565	455	110**	904	724	180**	1113	753	361**	2065	1772	293**	2523	2170	353**	2945	2364	581**
Andhra Pradesh	971	641	330**	829	827	2	1435	1085	350	2185	1837	348	2830	2529	300	3738	2455	1283**
Assam	137	126	12	365	290	75	284	466	-183**	732	707	25	1501	951	550**	1890	2156	-266
Bihar	404	341	62	734	474	260**	875	298	577**	1629	1460	170	2306	1659	647**	1547	1663	-117
Chhattisgarh	186	137	48**	324	275	49*	345	229	116	1854	1000	854**	1855	1263	592	4897	1736	3161**
Delhi	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1818	1333	486*	2333	2123	209	2251	2044	207
Gujarat	449	415	34	1006	700	306**	788	332	455**	2746	2111	635	2620	2509	112	3359	2362	996*
Haryana	1349	1108	241	1663	1651	12	2022	1182	840**	4835	4544	291	6047	4724	1323	7288	5539	1749
Himachal Pradesh	1770	1431	340*	2628	1963	665**	3604	2443	1161**	5368	4819	549	6084	5780	304	4961	3922	1039
Jammu & Kashmir	1147	1287	-140	2102	1926	176	2731	2293	438	3310	3078	232	5425	4091	1335	4882	3775	1106
Jharkhand	331	448	-117	719	551	168	416	925	-509	3164	2024	1139**	3559	3185	374	3958	2487	1471**
Karnataka	499	469	29	704	675	29	1120	739	381**	2903	2318	584*	2811	2260	551*	3147	2835	312
Kerala	1591	1355	236	2230	2083	146	3223	3623	-399	1356	1583	-227	2441	1986	455	2945	2210	735
Madhya Pradesh	294	267	27	531	369	162**	623	228	396**	1100	1135	-35	1651	1379	272	1306	1247	59
Maharashtra	193	207	-14	433	382	51*	700	401	299**	1321	1157	164	1712	1632	80	2046	1577	469*
Northeast	1445	1627	-182	1742	1556	186	1829	1444	386	3687	3980	-293	4158	4783	-625	5436	5176	260
Orissa	350	289	61	521	515	5	344	269	75	1807	1300	507*	2267	2159	108	1398	1994	-596
Punjab	1963	1614	348	2457	2007	450**	2319	2127	192	4649	4978	-329	5518	4445	1073**	4568	4839	-271
Rajasthan	622	431	191**	1065	615	450**	1137	442	695**	1735	1475	261	2441	1774	667**	2479	2189	290
Tamil Nadu	1076	468	608	738	748	-11	1839	1117	722	2052	1839	213	1720	1805	-85	3013	2307	706
Uttar Pradesh	405	289	116**	854	572	282**	1041	433	609**	1781	1491	290	2113	1984	128	2558	2061	497
Uttarakhand	820	937	-117	1354	992	362**	1309	1078	230	1952	3276	-1324*	4675	3986	689	5936	6318	-382
West Bengal	522	533	-11	1044	902	143	991	880	111	2412	2007	406	3410	2732	678	3828	3159	668

Note: [1]. Gender gap is the difference between average expenditure on male and female child in each age group. Positive gap implies pro-male bias, while negative gap implies pro-female bias. [2]. ** statistically significant at 5% level; *statistically significant at 10% level; N/A: not applicable.

Table3: Education Expenditure on enrolled children by age group, location and gender

State	Rural									Urban								
	Age 5-9			Age 10-14			Age 15-19			Age 5-9			Age 10-14			Age 15-19		
	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap	Male	Female	Gap
All India	740	621	119**	1045	902	143**	2302	2203	99	2424	2078	345**	2814	2491	323**	4961	4217	744**
Andhra Pradesh	1076	712	364**	943	1010	-67	2696	3519	-822	2310	1950	360	3118	2838	280	6229	4690	1539*
Assam	197	168	29	507	368	139	795	1094	-299*	844	856	-12	1702	1119	583*	2604	3049	-445
Bihar	685	625	60	910	759	151	2215	1133	1082**	2065	1807	257	2690	2069	621**	2746	3094	-348
Chhattisgarh	280	222	58*	367	331	36	748	1019	-272	2173	1111	1063**	2054	1515	538	6913	2778	4136*
Delhi	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	2236	1636	600*	2528	2304	224	3578	3424	153
Gujarat	532	536	-4	1122	859	263*	2029	1890	139	3309	2547	762	2929	3026	-98	6094	5448	646
Haryana	1651	1419	232	1852	1893	-41	3493	2744	748	5471	5007	464	7055	5291	1764	8722	8435	287
Himachal Pradesh	2007	1677	329	2726	2044	682**	4690	3286	1404**	5670	5005	666	6374	5948	426	5988	4767	1221
Jammu & Kashmir	1239	1430	-191	2301	2196	105	4406	4144	262	3499	3362	137	5621	4371	1251	5704	5191	513
Jharkhand	541	675	-134	911	863	48	1297	2640	-1343	3515	2599	917	3909	3571	338	5186	3874	1311*
Karnataka	663	648	15	810	775	35	2266	1977	290	3304	2660	644**	3051	2437	614**	5317	4991	326
Kerala	1962	1440	521*	2238	2083	154	4289	4932	-643	1712	2011	-299	2461	2003	458	3819	3269	550
Madhya Pradesh	414	410	4	610	462	147**	1209	798	411**	1328	1369	-41	1806	1658	149	2432	2418	14
Maharashtra	249	265	-15	468	436	32	1296	963	333**	1583	1401	183	1765	1747	18	3387	2595	792**
Northeast	1759	1917	-158	1886	1801	85	3151	2752	399	4279	4016	263	4385	4882	-497	6606	6120	486
Orissa	409	341	69	637	671	-34	1195	1485	-289	2016	1513	503	2621	2463	157	2945	4346	-1400
Punjab	2296	1872	424*	2704	2303	400*	4454	4386	68	5028	5203	-175	5988	4784	1204**	7377	6754	622
Rajasthan	799	662	136*	1231	907	324**	2412	1915	497*	2099	1842	257	2867	2253	614**	4511	4716	-204
Tamil Nadu	1250	589	661	770	811	-42	3003	2516	488	2285	1932	353	1817	1930	-113	4924	4286	638
Uttar Pradesh	522	390	132**	986	694	292**	2095	1376	719**	2198	1864	333	2716	2610	106	5489	4153	1336**
Uttarakhand	1033	1156	-123	1438	1096	342*	1801	1843	-42	3031	3667	-636	5295	4571	724	9003	9555	-553
West Bengal	647	699	-52	1343	1139	203	2649	2873	-224	2665	2362	303	3961	3363	598	6664	5705	959

Note: [1]. Gender gap is the difference between average expenditure on enrolled male and female child in each age group. Positive gap implies pro-male bias, while negative gap implies pro-female bias. [2]. ** statistically significant at 5% level; *statistically significant at 10% level; N/A: not applicable.

Table 4a: Difference in Marginal Effect (DME) × 100 of Gender Variables by age group (Household Results), All (Urban + Rural)

state	area	Age 5-9				Age 10-14				Age 15-19			
		Probit (1)	Conditiona IOLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona IOLS (4)	Probit (1)	Conditiona IOLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona IOLS (4)	Probit (1)	Conditiona IOLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona IOLS (4)
All India	ALL	6.18	0.93*	1.07**	0.78*	13.94**	0.7	1.24**	1.72**	17.49**	2.50**	2.92**	3.02**
Andhra Pradesh	ALL	6.88	1.53	1.63	2.86**	5.36	-0.79	-0.44	1.16	12	5.19**	5.04**	4.63*
Assam	ALL	6.1	-0.55	-0.26	-1.53	-49.95*	1.08	-0.38	0.64	24.76	-0.17	0.45	-0.59
Bihar	ALL	-1.46	-1.31	-1.08	-0.62	48.51**	4.33**	5.13**	5.95**	41.38**	3.9	4.54**	2.74
Chhattisgarh	ALL	4.81	4.15**	3.76**	2.58*	-1.39	1.58	1.32	2.74*	36.29**	4.13**	4.79**	4.52**
Delhi	ALL	6.75	-0.36	0.08	1.64	8.42	-3.62	-2.64	-0.36	5.33	-0.18	0.15	3.42
Gujarat	ALL	10.68	-4.05**	-2.5	-3.06**	2.35	0.58	0.51	0.54	32.46	5.47**	5.25**	2.90*
Haryana	ALL	3.1	3.06	2.8	1.94	-12.12	-2.26	-2.8	-4.07	10.16	2.45	2.82	0.55
Himachal Pradesh	ALL	-2.04	-2.15	-2.28	-1.36	-16.82**	0.74	-0.82	-1.48	0.37	8.56**	8.37**	8.18**
Jammu & Kashmir	ALL	-13.91	-0.61	-1.39	-0.74	79.87**	-3.81	1.22	-0.88	22.23	0.7	1.97	4.79
Jharkhand	ALL	12.52	4.71	4.07*	2.54	86.22**	0.36	4.25**	3.28	-13.38	-1.13	-1.45	-3.43
Karnataka	ALL	16.09	1.13	1.74	2.38*	-3.07	-0.86	-0.88	-0.79	15.01*	0.06	0.78	1.03
Kerala	ALL	21.42	5.00	6.27	2.74	-6.64	1.84	1.08	1.69	-2.5	0.32	0.07	-1.98
Madhya Pradesh	ALL	8.65	3.48**	3.25**	2.65**	22.47*	3.09**	3.44**	2.5	9.75	3.87**	3.62**	2.82
Maharashtra	ALL	-3.09	0.15	0.03	-0.62	5.68	0.23	0.38	1.22	17.43*	1.32	1.71*	3.10**
Northeast	ALL	-5.01	2.79	2.12	0.55	29.97	3.85	6.12**	3.25	11.32	1.22	2.1	3.24
Orissa	ALL	1.23	1.09	0.98	1.4	-0.67	-0.7	-0.63	-0.05	27.69**	4.45**	4.8	3.82**
Punjab	ALL	7.34	-1	-0.17	2.91	-11.63	9.46**	6.65**	6.85**	21.66	-2.79	-0.36	3.1
Rajasthan	ALL	24.21*	2.01	2.79	2.04	42.30**	4.35**	5.56**	4.61**	34.13**	6.62**	7.03**	5.86**
Tamil Nadu	ALL	-19.28	2.04	0.76	-0.26	3.61	-0.71	-0.45	2.83	-4.18	4.12	3.58	6.89
Uttar Pradesh	ALL	7.99	0.48	0.76	0.58	10.46	-1.03	-0.45	-0.82	11.68	2.74**	2.89**	1.75
Uttarakhand	ALL	10.80	-10.25*	-9.22	-8.27	0.30	5.56	5.49	4.07	12.48**	3.63	4.56	5.61
West Bengal	ALL	1.57	-1.42	-1.12	-0.61	14.64	-1.79	-0.63	1.2	21.64**	-1.61	-0.04	1.93

Note: [1]. ** statistically significant at 5% level; *statistically significant at 10% level. [2]. The cells display 100*DME; where DME is the difference in coefficients on proportion of male and female in that particular age group. [3]. Conditional OLS equation is fitted only for households with positive education spending; the dependent variable is natural log of household education budget share. The coefficient on the gender variable is transformed so that marginal effect reported in col. 2 are comparable to those in col. 4, where the dependent variable is in absolute rather than in log terms. Col. 4 pertains to the unconditional OLS of absolute household education budget share fitted to all households, including those with zero education budget shares. P-value for col. 3 is obtained by bootstrapping with 100 replications.

Table 4b: Difference in Marginal Effect (DME) × 100 of Gender Variables by age group (Household Results), Rural

State	area	Age 5-9				Age 10-14				Age 15-19			
		Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)	Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)	Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)
All India	Rural	8.96	0.86	1.06**	0.62	14.22**	0.68	1.12**	1.85**	23.31**	2.91**	3.31**	3.09**
Andhra Pradesh	Rural	14.70	0.59	1.05	1.70	-0.64	-1.28	-1.11	0.42	14.42	3.87*	3.81**	3.20
Assam	Rural	9.65	-0.26	0.03	-1.46	-67.25**	0.97	-0.80	1.35	22.36	-2.11	-0.93	-1.09
Bihar	Rural	-5.34	-1.50	-1.34	-0.68	55.21**	3.93**	4.88**	6.09**	42.20*	5.19*	5.43**	3.64
Chhattisgarh	Rural	16.84	4.66**	4.51**	3.19**	-2.77	1.37	1.11	2.66	38.83**	4.25**	4.77**	4.13**
Gujarat	Rural	11.42	-3.67*	-2.18	-2.98*	-4.79	0.06	-0.10	0.61	61.58**	6.60**	6.28**	2.93
Haryana	Rural	21.11	5.68*	6.15*	4.34*	-6.93	-1.36	-1.59	-2.79	9.53	2.48	2.71	-0.04
Himachal Pradesh	Rural	-4.62	-2.12	-2.47	-1.77	-15.67*	1.37	-0.04	-0.98	-0.82	8.52**	8.25**	8.88*
Jammu & Kashmir	Rural	-32.99	-1.02	-2.72	-4.99	88.97**	-4.68	0.36	-2.38	18.81	0.67	1.63	7.22*
Jharkhand	Rural	-3.30	4.52	2.96	1.43	116.94**	-0.65	4.30	4.24	-24.64	-0.52	-1.36	-4.57
Karnataka	Rural	12.53	0.02	0.51	0.90	-11.79	-0.70	-1.03	-0.90	19.39	0.51	1.17	0.65
Kerala	Rural	16.57	9.09	9.82*	5.18	-14.09	1.33	-0.04	1.88	-2.85	0.35	0.07	-4.86
Madhya Pradesh	Rural	18.64	2.61**	2.74**	2.79**	30.36*	3.36**	3.73**	4.15**	28.71**	4.94**	4.98**	5.95**
Maharashtra	Rural	-3.18	0.00	-0.06	-0.47	-5.20	0.05	-0.07	0.78	16.64	0.97	1.17	3.07**
Orissa	Rural	-0.19	1.09	0.93	1.36	-3.21	-0.63	-0.64	-0.21	39.90**	4.98**	5.51**	5.11**
Punjab	Rural	7.88	-0.20	0.41	3.09	-5.55	9.13**	7.58**	9.23**	15.94	-3.54	-1.91	1.19
Rajasthan	Rural	28.88*	2.46	3.14**	2.38*	57.75**	4.93**	6.29**	5.53**	48.24**	9.07**	9.22**	8.11**
Tamil Nadu	Rural	-14.08	9.44**	8.13	2.54	8.77	-0.18	0.27	5.11	-3.05	7.40	6.77**	10.78
Uttar Pradesh	Rural	13.27	0.12	0.57	0.08	3.76	0.10	0.22	0.63	15.68	3.67**	3.75**	3.72**
West Bengal	Rural	-1.34	-1.42	-1.27	-0.80	2.98	-2.63	-2.07	-1.15	32.81**	-3.27	-1.04	-0.63

Note: [1]. ** statistically significant at 5% level; *statistically significant at 10% level. [2]. The cells display 100*DME; where DME is the difference in coefficients on proportion of male and female in that particular age group. [3]. Conditional OLS equation is fitted only for households with positive education spending; the dependent variable is natural log of household education budget share. The coefficient on the gender variable is transformed so that marginal effect reported in col. 2 are comparable to those in col. 4, where the dependent variable is in absolute rather than in log terms. Col. 4 pertains to the unconditional OLS of absolute household education budget share fitted to all households, including those with zero education budget shares. P-value for col. 3 is obtained by bootstrapping with 100 replications. [4]. For Delhi only ALL is reported, as it is predominantly urban. In case of Northeast and Uttarakhand, only results from pooled sample reported because of restricted sample size.

Table 4c: Difference in Marginal Effect (DME) × 100 of Gender Variables by age group (Household Results), Urban

State	area	Age 5-9				Age 10-14				Age 15-19			
		Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)	Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)	Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)
All India	Urban	-0.45	1.24	1.05	1.26	12.97**	0.48	1.36	1.15	6.39*	1.06	1.40*	2.90**
Andhra Pradesh	Urban	-22.88	4.93	3.05	6.09**	15.41	2.67	3.62	3.31	3.42	8.84**	8.63**	8.56**
Assam	Urban	-13.00	-3.38	-3.67	-3.06	23.05	0.08	0.87	-0.77	18.87	2.18	2.72	0.73
Bihar	Urban	20.08	0.52	1.80	1.13	2.77	4.04	3.92	2.79	34.22**	-5.10	-2.46	-2.83
Chhattisgarh	Urban	-38.42	-0.25	-2.58	2.93	15.52	-2.63	-1.41	0.55	26.44	-0.55	1.13	5.81
Gujarat	Urban	3.89	-4.46	-3.26	-2.38	14.20	1.92	2.45	-0.09	5.77	3.47	3.11	2.25
Haryana	Urban	-27.33	-6.99	-9.44	-1.49	-41.03	-6.79	-10.84	-9.72	-7.29	3.78	2.58	-0.65
Himachal Pradesh	Urban	0.00	-2.41	-2.41	3.32	-0.00**	-10.99	-10.99	-7.78	0.00	8.08	8.08	6.34
Jammu & Kashmir	Urban	22.30	1.42	3.35	8.88	0.28	3.65	3.43	3.94	17.97	0.21	1.83	-0.65
Jharkhand	Urban	27.71	17.62**	20.02**	15.64**	-22.56	10.48	7.58	4.16	5.02	0.38	0.93	-1.60
Karnataka	Urban	18.83	5.74*	6.58**	4.69**	3.51	-2.69	-2.23	-1.46	3.86	-1.30	-0.93	1.68
Kerala	Urban	42.51*	-6.57	-2.64	-3.23	10.33	2.66	2.99	-0.03	10.17	1.64	2.11	9.22
Madhya Pradesh	Urban	-11.77	6.90	5.45	3.86	11.70	3.31	3.69	0.55	-32.04**	-2.87	-4.54	-4.60
Maharashtra	Urban	-2.08	-0.03	-0.14	-1.74	25.52	-0.10	1.32	1.89	20.62*	0.93	1.99	4.18
Orissa	Urban	8.17	2.50	2.81	2.24	11.86	-0.02	0.75	1.44	-10.79	0.47	-0.27	-2.58
Punjab	Urban	9.58	-7.66	-4.21	2.04	-25.94	6.55	1.52	-0.45	31.87	2.26	5.30	7.39
Rajasthan	Urban	3.95	3.70	3.54	1.61	4.35	0.83	1.06	1.90	-7.85	-4.18	-4.27	-2.73
Tamil Nadu	Urban	-23.67*	-3.51	-4.81	-1.48	-8.29	0.28	-0.28	2.08	-0.61	-0.22	-0.24	3.42
Uttar Pradesh	Urban	-15.08	5.76*	3.86	2.84	27.08**	-9.74**	-6.41*	-7.13**	8.47	0.63	1.29	-1.20
West Bengal	Urban	9.12	0.34	1.17	1.31	27.93**	1.30	3.83	6.19**	0.19	4.19	3.84	6.43*

Note: [1]. ** statistically significant at 5% level; *statistically significant at 10% level. [2]. The cells display 100*DME; where DME is the difference in coefficients on proportion of male and female in that particular age group. [3]. Conditional OLS equation is fitted only for households with positive education spending; the dependent variable is natural log of household education budget share. The coefficient on the gender variable is transformed so that marginal effect reported in col. 2 are comparable to those in col. 4, where the dependent variable is in absolute rather than in log terms. Col. 4 pertains to the unconditional OLS of absolute household education budget share fitted to all households, including those with zero education budget shares. P-value for col. 3 is obtained by bootstrapping with 100 replications. [4]. For Delhi only ALL is reported, as it is predominantly urban. In case of Northeast and Uttarakhand, only results from pooled sample reported because of restricted sample size.

Table 5a: Marginal Effect of the Male Dummy Variable (Individual-level data), All (Urban + Rural)

state	area	Age 5-9				Age 10-14				Age 15-19			
		Probit (1)	Conditiona IOLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona IOLS (4)	Probit (1)	Conditiona IOLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona IOLS (4)	Probit (1)	Conditiona IOLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona IOLS (4)
All India	ALL	0.02**	0.10**	0.09**	0.12**	0.05**	0.08**	0.10	0.13**	0.12**	0.07**	0.11**	0.13**
Andhra Pradesh	ALL	0.03	0.26**	0.24**	0.23**	0.04**	-0.02	0.02	0.04	0.18**	-0.13	0.11**	0.14**
Assam	ALL	0.00	0.00	0.00	0.06	-0.04	0.05	0.00	0.26**	-0.07	-0.11	-0.09*	-0.15
Bihar	ALL	0.03	0.10	0.08	0.03	0.15**	0.28**	0.33**	0.30**	0.13**	0.31**	0.19**	0.23**
Chhattisgarh	ALL	0.01	0.20**	0.14**	0.29**	0.06*	0.09	0.12**	0.21**	0.27**	-0.04	0.12**	0.32**
Delhi	ALL	0.01	0.03	0.03	0.25**	0.00	0.00	0.01	-0.02	0.01	0.06	0.04	0.05
Gujarat	ALL	0.04	-0.04	0.00	0.13	0.04	0.11	0.12**	0.10	0.17**	-0.09	0.10**	0.15**
Haryana	ALL	0.02	0.12**	0.11**	0.08	0.00	0.04	0.04	0.05	0.14**	0.10	0.16**	0.17**
Himachal Pradesh	ALL	0.01	0.04	0.04	0.04	-0.01	0.10**	0.09**	0.15**	0.01	0.20**	0.15**	0.28**
Jammu & Kashmir	ALL	0.03	0.03	0.05	-0.02	0.03	0.14	0.15	0.20**	0.09	0.06	0.12	0.10
Jharkhand	ALL	-0.01	0.1	0.05	0.18*	0.12**	0.07	0.11**	0.21**	0.04	0.05	0.03	-0.01
Karnataka	ALL	0.03	0.04	0.05	0.18**	0.01	0.05	0.05	0.07	0.08**	0.11*	0.11**	0.10**
Kerala	ALL	-0.05	0.16	0.07	0.07	0.00	0.08	0.09	0.13	0.00	-0.09	-0.06	-0.08
Madhya Pradesh	ALL	0.03	0.18**	0.14**	0.03	0.08**	0.24**	0.25**	0.21**	0.16**	0.22**	0.21**	0.18**
Maharashtra	ALL	-0.03	-0.03	-0.04	0.04	0.05**	0.01	0.04	0.04	0.08**	0.13**	0.13**	0.20**
Northeast	ALL	-0.05	0.14	0.08	0.13	0.02	-0.02	0.00	0.00	0.03	0.09	0.08	0.09
Orissa	ALL	0.04	0.10*	0.11**	0.17*	0.05	0.06	0.08**	0.07	0.07**	-0.20	0.02	0.00
Punjab	ALL	-0.02	0.01	0.00	-0.01	0.05*	0.15**	0.17**	0.22**	-0.03	0.04	-0.01	0.00
Rajasthan	ALL	0.11**	0.13**	0.18**	0.20**	0.15**	0.22**	0.30**	0.28**	0.20**	0.26**	0.23**	0.20**
Tamil Nadu	ALL	-0.02	0.11	0.08	0.14	0.01	-0.14**	-0.12**	-0.11	0.18**	-0.05	0.11**	0.18*
Uttar Pradesh	ALL	0.03	0.18**	0.16**	0.20**	0.03*	0.11**	0.12**	0.20**	0.15**	0.19**	0.17**	0.22**
Uttarakhand	ALL	-0.02	0.00	-0.01	-0.11	0.02	0.25**	0.26**	0.25**	0.12	0.07	0.13*	-0.03
West Bengal	ALL	0.05*	0.07	0.09**	0.07	0.00	0.04	0.04	0.13**	0.08**	0.02	0.05**	0.06*

Note: [1]. ** statistically significant at 5% level; *statistically significant at 10% level. [2]. The cells show the marginal effect of gender dummy variable (Male). In col. 2, 3, and 4, the marginal effect of gender dummy variable is divided by average expenditure per enrolled child in that state-area (Rural/Urban/ALL). [3]. For combined probit+ OLS models, the P-values were obtained from bootstrapping with 100 replications. [4]. The conditional OLS is fitted only for children with positive education spending, and the dependent variable is log of education spending. The coefficient of gender dummy variable were transformed so that the marginal effects reported in col. 2 are comparable to col. 4, where the dependent variable is in absolute term rather than in log terms. Col. 4 pertains to unconditional OLS of absolute education expenditure, fitted on all children, including those with zero education expenditure.

Table 5b: Marginal Effect of the Male Dummy Variable (Individual-level data), Rural

state	area	Age 5-9				Age 10-14				Age 15-19			
		Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)	Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)	Probit (1)	Conditional OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditional OLS (4)
All India	Rural	0.03**	0.13**	0.11	0.13**	0.06**	0.10**	0.13	0.17**	0.15**	0.08**	0.14**	0.15**
Andhra Pradesh	Rural	0.05	0.21**	0.21**	0.27**	0.06**	-0.06	0.00	0.04	0.22**	-0.32*	0.12**	0.11
Assam	Rural	0.00	0.03	0.02	0.07	-0.07	0.03	-0.04	0.15	-0.06	-0.20	-0.12	-0.14
Bihar	Rural	0.04	0.10	0.09**	0.04	0.17**	0.32**	0.38**	0.32**	0.14**	0.41**	0.22**	0.31**
Chhattisgarh	Rural	0.03	0.29**	0.21**	0.15*	0.05	0.11	0.15**	0.14**	0.28**	-0.28	0.20**	0.17*
Gujarat	Rural	0.05	-0.12	-0.04	0.01	0.06*	0.21**	0.23**	0.25**	0.19**	-0.34	0.14**	0.22**
Haryana	Rural	0.03	0.20**	0.18**	0.13	0.01	0.05	0.05	-0.01	0.13**	0.15	0.19**	0.23**
Himachal Pradesh	Rural	0.01	0.03	0.03**	0.04	-0.01	0.11**	0.10**	0.21**	0.00	0.20**	0.14**	0.28**
Jammu & Kashmir	Rural	0.02	0.05	0.06	0.00	0.03	0.13	0.15**	0.18	0.08	0.01	0.09**	0.09
Jharkhand	Rural	-0.03	0.17*	0.08**	0.10	0.12**	0.07	0.12**	0.25	0.02	0.03	0.02**	-0.05
Karnataka	Rural	0.02	-0.02	0.01	0.06	0.03	0.05	0.07**	0.04	0.11**	0.13	0.16**	0.18**
Kerala	Rural	-0.08	0.22	0.08	0.16	0.00	0.06	0.07	0.10	-0.03	-0.14	-0.12*	-0.15
Madhya Pradesh	Rural	0.05*	0.26**	0.21**	0.09	0.08**	0.31**	0.33**	0.32**	0.23**	0.30**	0.30**	0.37**
Maharashtra	Rural	-0.06	-0.08	-0.12**	-0.09	0.06**	0.01	0.07	0.10*	0.13**	0.16*	0.20**	0.29**
Orissa	Rural	0.05	0.09	0.11**	0.16	0.06*	0.05	0.09*	0.07	0.08**	-0.33	0.04**	0.07
Punjab	Rural	-0.02	0.03	0.01	0.08	0.07**	0.14*	0.19**	0.22**	0.01	0.01	0.01	0.02
Rajasthan	Rural	0.12**	0.13*	0.20**	0.26**	0.18**	0.29**	0.39**	0.37**	0.22**	0.39**	0.31**	0.33**
Tamil Nadu	Rural	0.03	0.26*	0.22	0.30	0.03	-0.19*	-0.14	-0.20	0.25*	-0.07	0.19	0.24
Uttar Pradesh	Rural	0.03	0.24**	0.21**	0.26**	0.04**	0.15**	0.16**	0.32**	0.19**	0.28**	0.24**	0.31**
West Bengal	Rural	0.04	0.13	0.14**	0.00	-0.01	0.04	0.02	0.10	0.08**	0.01	0.06**	0.03

Note: [1]. ** statistically significant at 5% level; *statistically significant at 10% level. [2]. The cells show the marginal effect of gender dummy variable (Male). In col. 2, 3, and 4, the marginal effect of gender dummy variable is divided by average expenditure per enrolled child in that state-area (Rural/Urban/ALL). [3]. For combined probit+ OLS models, the P-values were obtained from bootstrapping with 100 replications. [4]. The conditional OLS is fitted only for children with positive education spending, and the dependent variable is log of education spending. The coefficient of gender dummy variable were transformed so that the marginal effects reported in col. 2 are comparable to col. 4, where the dependent variable is in absolute term rather than in log terms. Col. 4 pertains to unconditional OLS of absolute education expenditure, fitted on all children, including those with zero education expenditure. [5]. For Delhi only ALL is reported, as it is predominantly urban. In case of Northeast, Uttarakhand, only results from pooled sample (ALL) is reported because of sample size considerations.

Table 5c: Marginal Effect of the Male Dummy Variable (Individual-level data), Urban

state	area	Age 5-9				Age 10-14				Age 15-19			
		Probit (1)	Conditiona I OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona I OLS (4)	Probit (1)	Conditiona I OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona I OLS (4)	Probit (1)	Conditiona I OLS (2)	Combined Probit+OLS (3)=f(1,2)	Unconditiona I OLS (4)
All India	Urban	0.00	0.11**	0.09**	0.12**	0.02**	0.06**	0.07**	0.08**	0.04**	0.07**	0.06**	0.11**
Andhra Pradesh	Urban	-0.01	0.49**	0.45**	0.20**	0.00	0.15*	0.14	0.04	0.04	0.15	0.10*	0.18**
Assam	Urban	0.08	-0.28	-0.18	-0.02	0.08	0.21	0.25	0.49**	-0.13*	0.03	-0.08	-0.23
Bihar	Urban	0.00	0.10	0.07	0.02	0.02	0.15*	0.14**	0.15	0.05	-0.08	0.02	0.02
Chhattisgarh	Urban	-0.11	0.47**	0.29	0.48**	0.10*	0.00	0.07	0.22	0.15	0.24	0.26**	0.57**
Gujarat	Urban	0.00	0.06	0.05	0.18	-0.02	-0.04	-0.05	-0.03	0.13**	-0.03	0.09*	0.15
Haryana	Urban	0.00	-0.04	-0.04	-0.03	-0.01	0.07	0.06	0.02	0.15	0.13	0.21*	0.17
Jharkhand	Urban	0.17**	0.09	0.20**	0.32**	0.06*	0.21*	0.24**	0.14	0.15**	0.19*	0.26**	0.25**
Kerala	Urban	0.02	0.05	0.05	-0.12	0.00	0.11	0.11	0.18	0.09	0.01	0.08	0.14
Karnataka	Urban	0.04	0.14*	0.15**	0.22**	-0.03	0.03	0.01	0.04	0.01	0.15*	0.09	0.05
Madhya Pradesh	Urban	-0.03	0.04	0.01	-0.05	0.04	0.17**	0.17**	0.10	-0.04	0.20*	0.05	-0.02
Maharashtra	Urban	0.01	0.00	0.01	0.07	0.03*	-0.01	0.02	0.00	0.03	0.13	0.10	0.20**
Orissa	Urban	0.01	0.27**	0.23**	0.27*	-0.02	0.13*	0.10	0.07	-0.05	0.05	-0.04	-0.17
Punjab	Urban	-0.03	-0.02	-0.04	-0.09	-0.01	0.22**	0.19**	0.20**	-0.10*	0.08	-0.03	-0.02
Rajasthan	Urban	0.04	0.19**	0.18**	0.15*	0.02	0.08	0.09	0.12*	0.10**	0.05	0.09**	-0.01
Tamil Nadu	Urban	-0.08**	0.01	-0.05	0.10	0.00	-0.03	-0.03	0.07	0.10**	-0.05	0.03	0.15
Uttar Pradesh	Urban	0.01	0.18**	0.15**	0.14*	-0.01	0.03	0.02	0.01	0.01	0.09	0.05	0.14**
West Bengal	Urban	0.09**	0.05	0.10	0.22	0.06*	0.06	0.10	0.22*	0.05	0.04	0.06	0.12**

Note: [1]. ** statistically significant at 5% level; *statistically significant at 10% level. [2]. The cells show the marginal effect of gender dummy variable (Male). In col. 2, 3, and 4, the marginal effect of gender dummy variable is divided by average expenditure per enrolled child in that state-area (Rural/Urban/ALL). [3]. For combined probit+ OLS models, the *P*-values were obtained from bootstrapping with 100 replications. [4]. The conditional OLS is fitted only for children with positive education spending, and the dependent variable is log of education spending. The coefficient of gender dummy variable were transformed so that the marginal effects reported in col. 2 are comparable to col. 4, where the dependent variable is in absolute term rather than in log terms. Col. 4 pertains to unconditional OLS of absolute education expenditure, fitted on all children, including those with zero education expenditure. [5]. For Delhi only ALL is reported, as it is predominantly urban. In case of Jammu & Kashmir, Himachal Pradesh, Northeast, Uttarakhand, results from pooled sample (ALL) is reported because of sample size considerations.

Table 6: Coefficient of the Male dummy in the individual level equations with family fixed effects

State	Age 5-9			Age 10-14			Age 15-19		
	ANYEDUEXP	EDUEXP (if>0)	EDUEXP	ANYEDUEXP	EDUEXP (if>0)	EDUEXP	ANYEDUEXP	EDUEXP (if>0)	EDUEXP
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
All India	0.04**	0.15**	0.14**	0.08**	0.15**	0.17**	0.12**	0.39**	0.32**
Andhra Pradesh	0.02	0.19	0.17	0.10**	0.11	0.15*	0.12**	-0.52*	-0.01
Assam	0.03	-0.06	0.01	0.08	-0.02	-0.01	0.04	-0.05	0.03
Bihar	0.15**	-0.07	0.28	0.20**	0.18	0.37**	0.05	0.07	0.39
Chhattisgarh	0.04	0.14	0.01	0.08	0.05	0.11	0.25**	4.78	0.93
Delhi	0.05	0.20	0.35**	0.02	-0.01	0.05	0.07	0.40**	0.33**
Gujarat	0.03	-0.07	0.01	0.13**	0.27	0.32**	0.15*	3.12*	0.52**
Haryana	0.06	0.10*	0.10*	0.04	0.14**	0.12**	0.17**	0.12	0.42**
Himachal Pradesh	0.00	0.11*	0.14*	-0.02	0.16	0.12	0.04	0.04	0.12
Jammu & Kashmir	-0.01	-0.06	-0.07	0.10**	0.11	0.21**	0.03	0.57*	0.40**
Jharkhand	-0.08	0.45**	0.09	0.03	0.38**	0.44**	0.16**	0.90**	0.51**
Karnataka	0.04	0.34**	0.21**	0.00	0.01	-0.05	0.08**	-0.32	0.09
Kerala	-0.17**	-0.05	-0.17	0.00	-0.01	-0.01	-0.02	1.76	0.68
Madhya Pradesh	0.07*	-0.04	0.03	0.11**	0.29**	0.29**	0.13**	0.36	0.24
Maharashtra	-0.03	0.11	0.02	0.05	-0.03	-0.01	0.11**	0.37	0.33**
Northeast	-0.12**	0.26**	0.22	0.02	0.01	-0.01	0.02	-0.06	-0.07
Orissa	0.01	0.17	0.12	0.10**	0.18*	0.18**	0.11	-2.58	0.54
Punjab	0.02	0.12*	0.02	0.07	0.05	0.04	-0.01	0.53	0.10
Rajasthan	0.10**	0.22**	0.18**	0.15**	0.30**	0.35**	0.17**	1.35**	0.71**
Tamil Nadu	-0.02	-0.06	-0.11	0.03	-0.05	-0.06	0.31**	-0.44	0.62**
Uttar Pradesh	0.06**	0.26**	0.30**	0.07**	0.31**	0.27**	0.13**	0.63*	0.35**
Uttarakhand	0.21**	0.51**	0.42**	0.00	0.13	0.12	0.11	-0.30	-0.21
West Bengal	0.06	0.07	0.21*	0.03	0.12	0.08	0.04	0.54*	0.26*

Note: [1]. ** statistically significant at 5% level; *statistically significant at 10% level. [2]. Three individual level equations for each age group are: (1) the probit estimation of ANYEDUEXP (whether any positive expenditure was incurred), (2) EDUEXP conditional on ANYEDUEXP==1, (3) unconditional education expenditure (zeros for those who did not spend on education). [3]. Marginal effects of Male dummy is reported. In col. 2 and 3, the marginal effect of gender dummy variable is divided by average expenditure per enrolled child in that state (sample include only those households which have both male and female children in that age group). [4]. Age of child is controlled.

Table 7: Gender difference in school choice, raw, LPM and fixed effect estimates

	Age 5-9				Age 10-14				Age 15-19			
	Private attendance	Gender difference (male-female) in private attendance	Coefficient on Male (OLS) with controls × 100	Coefficient on Male in fixed effect model × 100	Private attendance	Gender difference (male-female) in private attendance	Coefficient on Male (OLS) with controls × 100	Coefficient on Male in fixed effect model × 100	Private attendance	Gender difference (male-female) in private attendance	Coefficient on Male (OLS) with controls × 100	Coefficient on Male in fixed effect model × 100
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
All India	30.53	5.14**	5.19**	4.95**	24.32	3.85**	4.20**	5.45**	19.55	0.44	2.21**	2.86**
Andhra Pradesh	40.71	9.83**	10.03**	9.07**	24.27	5.39*	6.64**	5.22**	14.24	0.40	6.25*	3.49
Assam	5.78	1.85	2.50	0.00	6.19	0.00	-0.47	-1.76	16.67	-4.60	-2.46	0.00
Bihar	20.75	2.21	1.95	3.38	15.32	-2.37	1.18	3.84	9.81	-4.03	-0.31	7.08
Chhattisgarh	21.85	4.72	5.96*	6.92	12.00	6.57**	5.88**	6.07*	17.24	-6.05	-5.74	0.11
Delhi	34.09	2.18	1.34	1.18	24.96	6.62*	3.62	6.27**	20.82	4.47	6.75	8.02
Gujarat	22.74	1.70	0.34	-0.13	22.75	-0.63	0.07	0.74	23.08	-5.99	-0.69	7.86
Haryana	56.81	9.97**	7.12**	10.00**	39.41	6.27	4.86	9.05**	37.51	5.04	9.06	0.64
Himachal Pradesh	26.47	8.49**	2.54	4.70	15.53	8.03**	4.78*	3.09	9.90	6.20**	6.66**	-1.88
Jammu & Kashmir	45.98	1.34	3.64	-9.76*	44.68	4.92	10.19**	15.74**	17.35	7.76*	4.70	6.30
Jharkhand	33.61	-0.37	3.05	8.86**	32.26	1.22	2.59	12.77**	24.85	-0.87	-1.47	13.14
Karnataka	33.20	0.88	2.50	7.49**	25.02	3.89	3.56*	4.31**	23.24	8.09**	8.58**	2.04
Kerala	40.43	2.25	-0.85	-0.53	23.07	-4.98	-5.18	-1.92	23.19	-1.34	-2.55	-26.05**
Madhya Pradesh	31.34	5.67**	7.10**	9.51**	25.29	8.13**	8.55**	6.19**	25.90	0.99	8.32**	12.09**
Maharashtra	18.18	2.58	0.18	2.86	19.36	2.55	2.80	1.64	17.43	-4.39	-1.24	0.67
Northeast	42.31	-2.70	4.47	9.21**	29.32	-3.54	-1.30	1.03	23.05	5.49	6.36	-12.15*
Orissa	7.80	1.52	2.00	0.65	8.15	-0.50	0.08	2.06	15.46	-19.69**	-19.59**	-22.38
Punjab	61.20	5.59	5.14	4.68**	47.96	4.29	7.16*	7.68**	36.10	6.72	6.29	8.50
Rajasthan	35.87	4.37	7.66**	6.12**	29.73	11.10**	11.16**	8.57**	20.72	-1.42	2.63	8.61
Tamil Nadu	33.21	6.28	6.10*	1.53	17.53	-5.10	-3.49	3.99	15.16	-4.10	-1.10	-6.64
Uttar Pradesh	38.12	9.93**	10.43**	4.10**	38.11	8.93**	7.59**	10.38**	24.01	2.20	4.33	1.18
Uttarakhand	33.98	0.16	4.19	19.47*	21.80	15.90**	12.89**	6.88	19.19	21.24**	19.64**	33.64**
West Bengal	14.43	1.46	1.45	4.99*	5.81	2.76**	1.65	-0.14	5.47	1.13	1.11	-1.03

Note: [1]. ** statistically significant at 5% level; *statistically significant at 10% level. [2]. Private attendance refers to percentage of total enrolled children attending private school; col. 2 report gender difference in private attendance among enrolled children. [3]. Col. 3 reports the coefficient on Male dummy from a LPM model fitted in sample of all enrolled children. The controls include age of child, log of monthly per capita (lpce), square of lpce, log of household size, head years of education, and dummies for male, head's occupation, urban, Scheduled Caste, Scheduled Tribe, Other Backward Castes, and Muslim. [4]. Col. 4 reports the coefficient on Male dummy from LPM model with household fixed effects fitted in the sample of only those households which have both male and female children enrolled in that age group. Age of child is controlled.

Appendix

Table A1: OLS regression of budget share of education; binary probit of any education expenditure, and OLS regression on natural log of budget share of education

Variable	Description	Unconditional OLS (EDUSHARE)	Probit Marginal-effects (ANYEDUEXP)	Conditional OLS (LNEDUSHARE)
		(1)	(2)	(3)
LNPCE	Log of per capita expenditure	0.01	0.31***	1.02***
LNPCESQ	Square of LNPCE	0.00	-0.02***	-0.08***
LNHHSIZE	Log of household size	0.02***	0.25***	0.06*
M0to4	Fraction of household's members in			
	----- male & 0-4 age group	-0.05***	-0.24***	-1.88***
M5to9	----- male & 5-9 age group	0.02**	0.64***	-0.30*
M10to14	----- male & 10-14 age group	0.05***	0.80***	0.56***
M15to19	----- male & 15-19 age group	0.06***	0.09	1.41***
M20to24	----- male & 20-14 age group	-0.03***	-0.37***	-0.26
M25to60	----- male & 25-60 age group	-0.03***	-0.24***	-0.56***
M61plus	----- male & above 61 age	-0.03***	-0.20**	-0.10
F0to4	----- female & 0-4 age group	-0.05***	-0.24***	-1.59***
F5to9	----- female & 5-9 age group	0.01	0.58***	-0.50***
F10to14	----- female & 10-14 age group	0.03***	0.66***	0.41**
F15to19	----- female & 15-19 age group	0.03***	-0.09	0.88***
F20to24	----- female & 20-24 age group	-0.01*	-0.15**	-0.58***
F25to60	----- female & 25-60 age group	-0.00	0.16***	-0.07
FEMALEHEAD	Family head is female	0.00*	-0.01	0.14***
HEADYRSEDN	Years of education of head	0.00***	0.01***	0.05***
WAGELABOR	Household head is wage laborer	-0.01***	-0.04***	-0.19***
SC	Indicator for Scheduled Castes	-0.01***	-0.03***	-0.20***
ST	Indicator for Scheduled Tribes	-0.01***	-0.04**	-0.22***
OBC	Indicator for Other Backward Castes	-0.01***	-0.02**	-0.10***
MUSLIM	Indicator for Muslim	-0.02***	-0.11***	-0.29***
URBAN	Urban residence	0.02***	-0.00	0.54***
Constant		-0.04		
Observations		30199	30199	23752
R-squared		0.21		0.29
P-Values (M=F):				
Age 5-9		0.06	0.14	0.08
Age 10-14		0.00	0.00	0.16
Age 15-19		0.00	0.00	0.00

Note: *** p<0.01, ** p<0.05, * p<0.1. The model also includes state dummies not reported in table.

Table A2: Annual average education expenditure (Indian Rupees) on enrolled child, by type of school

State	Age 5-9		Age 10-14		Age 15-19	
	Public	Private	Public	Private	Public	Private
All India	468	2694	880	3170	3027	4747
Andhra Pradesh	347	2604	705	3949	3861	5432
Assam	220	1082	544	1996	1522	2931
Bihar	538	2205	797	2673	2149	3565
Chhattisgarh	235	1830	370	2054	1988	3521
Delhi	845	4363	1165	6269	2859	6739
Gujarat	537	4516	980	4050	3766	7293
Haryana	838	3759	1161	4946	4465	6603
Himachal Pradesh	1246	5121	1978	7321	3847	9098
Jammu & Kashmir	594	2966	1435	4175	4483	6099
Jharkhand	408	2367	621	3405	2677	3708
Karnataka	510	4034	751	3810	4043	4123
Kerala	1238	2834	1705	3826	4494	5017
Madhya Pradesh	254	1622	390	2234	1224	2926
Maharashtra, Goa	464	2675	718	2249	2231	3372
Northeast	1003	4286	1713	4150	3716	4946
Orissa	377	3028	815	2665	2028	2558
Punjab	962	4628	1716	5501	5764	6557
Rajasthan	466	2108	830	3004	2309	6003
Tamil Nadu	460	4025	687	3647	3388	5579
Uttar Pradesh	304	1508	792	1823	2630	3089
Uttarakhand	704	2928	1145	4182	2556	6299
West Bengal	619	4186	1488	6827	3706	11877

Note: Education expenditure is derived by adding school fees, expenditures on private tuition, books, uniforms, transportation and other materials.