# Intrahousehold Gender Gap in Education Expenditure in Bangladesh*

Sijia Xu[†], Abu S. Shonchoy[‡], and Tomoki Fujii[§]

March 15, 2018

*We are very grateful to IDE-JETRO for providing us with funding for this research. We would like to thank Niaz Asadullah, Yvonne Jie Chen, Namrata Chindarkar, Hai-Anh H. Dang, Isaac Ehrlich, Nobu Fuwa, Tatsuo Hatta, Christine Ho, Charles Yuji Horioka, Ravi Kanbur, Jong-Wha Lee, Peng Liu, Sunha Myong, Manabu Nose, Yasuyuki Sawada, Yoshito Takasaki, Long Q. Trinh, Wuyi Wang, Naoyuki Yoshino, and participants at seminars and workshops organized by the University of Tokyo, the National University of Singapore, and Asia Growth Research Institute and Asian Development Bank Institute for their comments and inputs.

[†]Singapore Management University (email: sijia.xu.2013@phdecons.smu.edu.sg)
[‡]Institute of Developing Economies (IDE) - JETRO and New York University (email: parves.shonchoy@gmail.com)
[§]Singapore Management University (email: tfujii@smu.edu.sg)

# Intrahousehold Gender Gap in Education Expenditure in Bangladesh

**Abstract**

Bangladesh has witnessed a reversal of gender gap in enrollment, from pro-male to pro-female, in the past decades. Nevertheless, the education outcomes for girls appear to have consistently lagged behind that for boys. We investigate this issue by elucidating the gender gap in intrahousehold allocation of education resources with a three-part model, which decomposes the households' education decisions into the following three parts: enrollment, conditional education expenditure, and share of education expenditure allocated to the core educational items, or items directly related to the quality of education. The model further incorporates the possible interdependence across these three decisions. Using four rounds of the Household Income and Expenditure Survey data, we find a pro-female bias in enrollment decision but a pro-male bias in the decisions on the conditional expenditure and core share in education expenditure from 2000 onwards. This apparent inconsistency of gender bias seems to be partly driven by the Female Stipend Programs (FSPs). FSPs have played an important role in promoting girls' enrollment in secondary schools but did not help to close the gender gap in conditional expenditure and core share allocation. Furthermore, the FSPs did not help narrow the gender gap in timely graduation from secondary school among primary-school graduates. Taken together, our empirical evidence suggests that the gender gap in the investment in the quality of education persisted in Bangladesh.

**JEL Classification**: D15, I28, J16, O15

**Keywords**: gender gap; education; female stipend program; hurdle model; Bangladesh

## 1 Introduction

Bangladesh has made a remarkable progress in gender equality in education over the past two decades. Intensive education investment and interventions, particularly in girls, helped narrow the gender gap in the school enrollment, highest grade attained, and some other educational indicators (Ahmed et al., 2007). According to BANBEIS (2006), only 34 percent of students enrolled in secondary schools were girls in 1990, but this figure exceeded half by 1998. This success in closing the gender gap in secondary school enrollment has indeed attracted much attention from researchers. Various studies (Asadullah and Chaudhury, 2009; Behrman, 2015; Khandker et al., 2003; Mahmud, 2003) indicate that the success in closing the gender gap in secondary school enrollment owe at least partly to the stipend and tuition fee waiver targeted at girls through various programs by the Government of Bangladesh (GOB) and donor agencies, which we collectively refer to as the Female Stipend Programs (FSPs). Begum et al. (2017) further show that the FSPs also benefit the siblings of the affected children, creating indirect, long-term gains for the society.

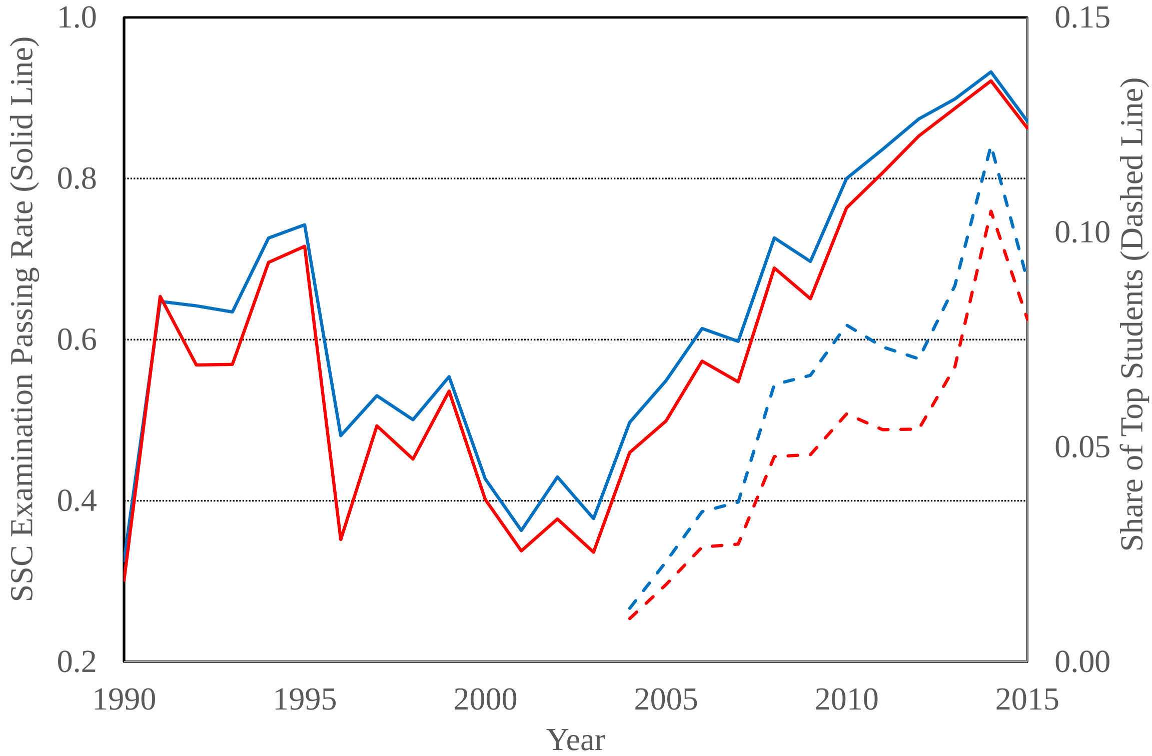


Figure 1: The solid lines represent the proportion of boys (blue) and girls (red) who have passed the Secondary School Certificate (SSC) examination among those who took the exam and the dashed lines represent the share of top students who achieved the highest grade point average (GPA 5). *Source: BANBEIS-Education Database* (`http://data.banbeis.gov.bd/`) accessed on Oct 29, 2017.

Despite this improvement, girls lagged behind boys in the education outcomes at the secondary level. Girls consistently underperformed boys both in terms of the passing rate of the Secondary School Certificate (SSC) examination and the share of top students who achieved the highest grade point average (GPA 5) in the SSC exam as Figure 1 shows. Girls are also found to have higher rates of dropout and grade repetition (Schurmann, 2009).

These observations appear to suggest that the investment in the quality of education may have been lower for girls than for boys, leading to the girls' relative underperformance in education. We, therefore, study the gender gap in the allocation of educational expenditure to investigate the possibility that the quality of education for girls may have been poorer than that for boys conditional on enrollment. These observations appear to suggest that the investment in the quality of education may have been lower for girls than for boys, leading to the girls' relative underperformance in education.

To this end, we develop a three-part model consisting of the following three related decisions on the education of children in a household: 1) enrollment,[1] 2) amount of education expenditure condi-

[1] We define enrollment to be one if the education expenditure is positive and the child is enrolled in secondary school, and zero otherwise. Around 0.39% of observations reporting to have enrolled in secondary school with zero education expenditure are dropped. Thus, enrollment here refers to secondary school enrollment with positive educational expenditure for the secondary-school age group.

tional on enrollment, and 3) share of education expenditure allocated to the core component, which directly relates to the quality of education as elaborated subsequently. Our model can be viewed as an extension of the hurdle model adopted by Kingdon (2005), which only includes the first two decisions, to incorporate a separate decision making for the investment in the quality of education. Therefore, unlike Kingdon (2005), we are able to detect the gender difference in the share of education expenditure allocated to the core component, even if the total education expenditure is the same between boys and girls.

Our three-part model has three noteworthy features. First, as with Kingdon (2005), our model separates the parental decision on the investment in education into the extensive and intensive margins—whether the child is enrolled in secondary school and how much is spent on education conditional on enrollment. This separation is important particularly when analyzing the gender gap, because school enrollment only reflects the quantity of education but not quality. Put differently, the education investment in girls conditional on enrollment may be lower than that in boys, even when the girls has a higher enrollment rate than boys.

Second, unlike Kingdon (2005), our model allows us to account for the gender difference in how the education expenditure is used, a point that is mostly neglected in the literature. To see the relevance of this point, consider a household with a boy and a girl in which an equal amount is spent on the education of each child. Suppose further that the education expenditure for the boy is mostly used to pay for home tutoring whereas that for the girl is mostly used to buy better or more uniforms. This gender difference in the pattern of education expenditure would reflect the gender difference in the quality of education that they receive.

Third, our three-part model takes into account the correlations of the three decisions conditional on observable characteristics. This is important because there may be some unobservable characteristics, such as innate ability, which may affect all three decisions simultaneously. For example, a smart child is more likely to be enrolled in school due to the higher expected returns from education. However, the child may require less education expenditure from the household than a less smart counterpart, because of a lower need for home tutoring or higher chance of receiving merit-based scholarships, for example. On the other hand, households may be more encouraged to invest in children with a higher ability to learn.

We apply the three-part model to the observations of school-age children from a total of four rounds of household surveys. Our analysis indicates that there is a pro-female bias in the enrollment decision, but the decisions on the total education expenditure and core share conditional on enrollment are biased against female in recent rounds. While this gap exists both at the primary and secondary levels, it is much more pronounced at the secondary level.

Our analysis also shows that the pro-female bias in enrollment became stronger between 1995 and 2010. On the other hand, the strong pro-male bias in conditional expenditure did not change much at the secondary level. Further, the decision on the core share allocation has become more pro-male. This finding is interesting because such inconsistency in the direction of gender bias is unique to Bangladesh to the best of our knowledge. In particular, existing studies in other South Asian countries such as India and Pakistan tend to find pro-male bias as elaborated in the next section.

Therefore, a natural question that arises here is why the parents in Bangladesh behave differently from other south Asian countries that share the historical roots and have broadly similar cultural, political, and economic backgrounds. Clearly, gender discrimination alone fails to explain what is observed in Bangladesh, because it would also lead to pro-male bias in enrollment. We, therefore, explore the relevance of the stipend program, because a comparable nationwide program does not exist in India or Pakistan. We indeed find some evidence that the FSPs help explain the inconsistency in the direction of bias. Therefore, while a program like FSPs may help improve or even reverse the gender gap in the quantity of education, it does not necessarily fill the gap in the quality of education. Hence, even though policies to narrow the gender gap in the quantity of education are desirable, policy-makers should be also wary of the potential implications for the quality of education.

The rest of this paper is organized as follows. Section 2 reviews related studies and discusses our paper's relevance and contributions to the existing studies. Section 3 introduces the three-part model. Section 4 describes the data and reports key summary statistics. Main empirical findings are presented in Section 5. Section 6 investigates the relevance of FSPs to observed pattern of gender bias pattern. We then provide a diagrammatical analysis in Section 7 to explain our findings, followed by the conclusion in Section 8.

## 2 Relevance to Existing Studies and our Contributions

Classical household theory suggests that decisions on intrahousehold resource allocation depend on preferences, investment returns, and time and income constraints (Behrman et al., 1982). Preferences may change over time with changing social norms. For example, Blunch and Das (2015) reports that younger cohorts have a more positive attitude towards gender equality than older cohorts in Bangladesh. Invest returns also matter. As Asadullah (2006) argues, if the labor market returns to education for girls is higher than that for boys, parents would be more motivated to invest in the former. On the other hand, if education of girls is deemed to bring about no returns to their parents or to lower the prospect of marriage, parents may be discouraged to invest in girls. This argument is true even when parents have no inherent gender bias. Indeed, this possibility is consistent with the experiment by Begum et al. (2016), who find that no systematic inherent gender bias by parents. Therefore, as noted by Lehmann et al. (2012), human capital investment in children in the same household may vary by a number of factors such as cognitive endowments, gender, and age and income of parents at the time of birth even though children possess similar genetic endowments. The current study relates to these studies by not only examining how much is spend on each child's education but also by dissecting the way the education resources are spent.

Many studies have found that parents tend to invest systematically more in sons than in daughters in developing countries (e.g., Deaton (1989) and Li and Tsang (2003)). This study relates to this literature and is built in particular on Kingdon (2005), who first incorporated the Working-Leser specification of Engel Curve approach into the hurdle model to study the gender bias in education expenditure in rural India. This model allows for the following two separate channels through which gender bias may exhibit: (1) enrollment and (2) conditional education expenditure. She found a pro-male bias in the enrollment decision but found no evidence of gender bias in educational expenditure among enrolled children. Azam and Kingdon (2013) revisit this study with more comprehensive data from India and find that the pro-male bias persists. This finding is also supported by Majumder et al. (2016) using Heckman's two-step model in West Bengal and Saha (2013) using the Oaxaca-Blinder decomposition approach.

Besides India, the hurdle models have been applied to other countries, including Kenayathulla (2016) in Malaysia, Aslam and Kingdon (2008) in Pakistan, Masterson (2012) in Paraguay, and Himaz (2010) in Sri Lanka as summarized in Table 2. This table shows that pro-male bias is not ubiquitous; pro-female bias was detected in Sri Lanka and no gender bias was found in Malaysia. Wongmonta and Glewwe (2017) also find a bias in favor of females in Thailand, though they do not use a hurdle model.

Table 2 also shows that the directions of the biases for enrollment and conditional education expenditure are always consistent (i.e., if one of them is positive and significant, then the other is never negative and significant). Furthermore, the direction of the bias was pro-male both in India (Azam and Kingdon, 2013; Kingdon, 2005) and Pakistan (Aslam and Kingdon, 2008), which share historical, socioeconomic, and cultural (India) or religious (Pakistan) backgrounds with Bangladesh. In contrast, we find an inconsistency in the direction of gender bias with a pro-female bias in enrollment and pro-male bias in total educational expenditure as elaborated subsequently.

Table 1: Existing Studies Using Hurdle Model

| **Paper** | **Data & Year** | **Sample Age** | $d$ | **Cond** $y$ |
|---|---|---|---|---|
| Kingdon (2005) | 16 states in Rural India, 1994 | 5 to 14 | $-$ | $\approx$ |
| Aslam and Kingdon (2008) | Pakistan, 2001-2002 | 5 to 9 | $-$ | $\approx$ |
| | | 10 to 14 | $-$ | $-$ |
| Himaz (2010) | Sri Lanka, 1990-91, 1995-96, 2000-01 | 5 to 9 | $\approx$ | $+$ |
| | | 10 to 13 | $\approx$ | $\approx$ |
| | | 14 to 16 | $\approx$ | $+$ |
| Masterson (2012) | Paraguay, 2000-2001 | 5 to 14 (Rural) | $-$ | $-$ |
| | | 5 to 14 (Urban) | $+$ | $+$ |
| Azam and Kingdon (2013) | India, 2004-05 | 5 to 9 | $\approx$ | $-$ |
| | | 10 to 14 | $-$ | $-$ |
| Kenayathulla (2016) | Malaysia, 2004-05 | 5 to 14 | $\approx$ | $\approx$ |

Note: $-$, $+$ and $\approx$ mean pro-male bias, pro-female bias and no bias, respectively.

This study makes contributions to the following three broad areas. First, the inconsistency in the direction of biases found in this study is new. It is also interesting because no such inconsistency was found in other countries including India and Pakistan as we have seen above.

Second, we make a modest but relevant contribution to the body of literature on limited dependent variable models. Our model is related to the double hurdle model originally proposed by Cragg (1971), which has been further extended and applied to study, among others, the consumption of food away from home (Yen, 1993), tobacco consumption (Aristei and Pieroni, 2008), and beef consumption (Jones and Yen, 2000) besides the studies on education expenditure mentioned above. Both the double-hurdle model and ours deal with a situation where the consumption amount is observed only when certain conditions are satisfied. However, ours is different because it has a third equation to model the core share in the total education expenditure. Furthermore, we allow for possible correlations in the unobservable error terms across different decisions. By taking advantage of this correlational structure, we are potentially able to obtain more accurate coefficient estimates than equation-by-equation regressions. The flexibility of our model enables us to detect the inconsistency in the direction of gender bias.

Finally, we provide a simple theoretical model somewhat similar to Dang and Rogers (2015) to explain why the policies like the FSPs may translate into the narrowing the gender gap in enrollment but not in the quality of education. Our theoretical model also provides a plausible explanation on why girls have underperformed boys in the SSC exam and other education outcomes. It also provides a cautionary lesson to researchers and policymakers that just increasing the enrollment of female students does not automatically lead to a greater gender equality in the quality of education children receive.[2]

## 3 The Three-Part Model

We extend the standard hurdle model, which consists of decisions on the school enrollment of the child and the amount of education expenditure conditional on enrollment, in two ways. First, we allow for correlations in the unobservable error terms across all the equations. Second, we incorporate the share of education expenditure on the core component in our model as the third part, which allows us to analyze the way education expenditure is spent. The education expenditure not spent on the core component is spent on the peripheral component, which do not directly relate to the quality of education. The definitions of the core and peripheral components are given in the next section.

Our three-part model formally has the following structure:

- Enrollment decision ($d$)

$$d = \mathbf{1}(x_d'\beta_d + \epsilon_d > 0), \tag{1}$$

  where $x_d$, $\beta_d$, and $\epsilon_d$ are covariates, their coefficient vector, and idiosyncratic error term for the enrollment equation, respectively, where the covariates include a dummy variable for girl to identify the gender effect. We use the subscript $d$ to denote the decision on $d$. We also use the subscripts $y$ and $s$ below in a similar manner.

- Education expenditure decision ($y$)

$$\log(y) = x_y'\beta_y + \epsilon_y \tag{2}$$

- Core component share decision ($s$)

$$s = \begin{cases} 0 & s^* \leq 0 \\ s^* & 0 < s^* < 1 \\ 1 & s^* \geq 1 \end{cases} \tag{3}$$

  where $s^* = x_s'\beta_s + \epsilon_s$ is the latent variable for $s$.

Note that education expenditure ($y$) and core component share ($s$) are observed only when the child is enrolled in school (i.e., $d = 1$).

[2] A related point was made in Shonchoy and Rabbani (2015). However, we provide more complete explanations of this phenomenon with a theoretical model and more rounds of data. We also investigate the gender differences in educational performance.

To allow for the dependency across the three equations, we assume that the error terms $\epsilon_d$, $\epsilon_y$, and $\epsilon_s$ have the following trivariate normal distribution:

$$\begin{bmatrix} \epsilon_d \\ \epsilon_y \\ \epsilon_s \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} 1 & \rho_{dy}\sigma_y & \rho_{ds}\sigma_s \\ \rho_{dy}\sigma_y & \sigma_y^2 & \rho_{ys}\sigma_y\sigma_s \\ \rho_{ds}\sigma_s & \rho_{ys}\sigma_y\sigma_s & \sigma_s^2 \end{bmatrix}\right), \tag{4}$$

where the variance of $\epsilon_d$ can be assumed to be unity without loss of generality.

There are four cases to consider in this setup: 1) the child is not enrolled in school ($d = 0$), 2) the child is enrolled in school with all education expenditure going to the peripheral component ($d = 1$ and $s = 0$), 3) the child is enrolled in school with education expenditure going to both the core and peripheral components ($d = 1$ and $0 < s < 1$), and 4) the child is enrolled in school with all education expenditure going to the core component ($d = 1$ and $s = 1$).[3]

Given the model structure described by eqs. (1)-(4), the log-likelihood $l_i$ for child $i$ given the parameter vector $\theta \equiv (\beta_d, \beta_y, \beta_s, \sigma_y, \sigma_s, \rho_{dy}, \rho_{ds}, \rho_{ys})^T$ can be written as follows:[4]

$$\begin{aligned} l_i(\theta) &= \mathbf{1}[d_i = 0] \cdot l_i^1 + \mathbf{1}[d_i = 1, y_i = y, s_i = 0] \cdot l_i^2 \\ &\quad + \mathbf{1}[d_i = 1, y_i = y, 0 < s_i < 1] \cdot l_i^3 + \mathbf{1}[d_i = 1, y_i = y, s_i = 1] \cdot l_i^4, \end{aligned}$$

where the log-likelihood $l_i^j$ for case $j \in \{1, 2, 3, 4\}$ is given by the following with $e_y \equiv \frac{\log(y) - x_y'\beta_y}{\sigma_y}$ and $e_s \equiv \frac{s - x_s'\beta_s}{\sigma_s}$:

$$\begin{cases} l_i^1 = \log\left[\Phi(-x'_{d_i}\beta_d)\right] \\ l_i^2 = \log(\phi(e_{y_i})) - \log(y_i) - \log(\sigma_y) \\ \qquad + \log\left[\Psi\left(\frac{x'_{d_i}\beta_d + \rho_{dy}e_{y_i}}{\sqrt{1-\rho_{dy}^2}}, -\frac{x'_{s_i}\beta_s + \rho_{ys}\sigma_s e_{y_i}}{\sigma_s\sqrt{1-\rho_{ys}^2}}, \frac{\rho_{dy}\rho_{ys} - \rho_{ds}}{\sqrt{(1-\rho_{dy}^2)(1-\rho_{ys}^2)}}\right)\right] \\ l_i^3 = \log\left(\phi\left(\frac{e_{y_i}}{\sqrt{1-\rho_{ys}^2}}\right)\right) + \log\left(\phi\left(\frac{e_{s_i}}{\sqrt{1-\rho_{ys}^2}}\right)\right) \\ \qquad + \left(\rho_{ys}\frac{e_{y_i}e_{s_i}}{1-\rho_{ys}^2}\right) - \log(y_i) - \log(\sigma_y) - \log(\sigma_s) - \log(\sqrt{1-\rho_{ys}^2}) \\ \qquad + \log\left[\Phi\left(\frac{x'_{d_i}\beta_d(1-\rho_{ys}^2) + (\rho_{dy} - \rho_{ds}\rho_{ys})e_{y_i} + (\rho_{ds} - \rho_{dy}\rho_{ys})e_{s_i}}{\sqrt{(1-\rho_{ys}^2 - \rho_{dy}^2 - \rho_{ds}^2 + 2\rho_{dy}\rho_{ds}\rho_{ys})(1-\rho_{ys}^2)}}\right)\right] \\ l_i^4 = \log(\phi(e_{y_i})) - \log(y_i) - \log(\sigma_y) \\ \qquad + \log\left[\Psi\left(\frac{x'_{d_i}\beta_d + \rho_{dy}e_{y_i}}{\sqrt{1-\rho_{dy}^2}}, \frac{x'_{s_i}\beta_s - 1 + \rho_{ys}\sigma_s e_{y_i}}{\sigma_s\sqrt{1-\rho_{ys}^2}}, \frac{\rho_{ds} - \rho_{dy}\rho_{ys}}{\sqrt{(1-\rho_{dy}^2)(1-\rho_{ys}^2)}}\right)\right]. \end{cases}$$

[3] Cases 2) and 4) are relatively rare in our data, accounting for 0.27% and 0.22% of all observations across years, respectively.

[4] The detailed derivation of the likelihood function for each cases is provided in Appendix A.

The sample log-likelihood function is just the summation of individual log likelihood function. Therefore, the maximum-likelihood (ML) estimator $\hat{\theta}_{ML}$ for the three-part model can be obtained as follows:

$$\hat{\theta}_{ML} = \arg\max_{\theta} \sum_{i=1}^{N} l_i(\theta).$$

The primary coefficients of interest are those on the girl dummy in $\beta_d$, $\beta_y$, and $\beta_s$. If the signs on these coefficients are positive, they indicate a pro-female bias but the opposite is true if the sign is negative. It should be noted here that the size of the coefficient does not necessarily equate with the size of the effect, because the model is nonlinear. Therefore, using the ML estimates, we calculate the marginal effects of being a girl on the probability of enrollment as well as conditional and unconditional levels of the total education expenditure and core expenditure. Because we cannot obtain a simple closed-form solution for the marginal effect due to the correlation across error terms, we need to use numerical integration to calculate marginal effects. The effects of girl on these quantities are computed as the change in the expected value of the outcome of interest when the value of the girl dummy variable changes from zero to one, where we use the following expressions for the conditional and unconditional expectations:

$$\begin{aligned}
E(d) &= P(d=1) = \Phi\left(x_d'\beta_1\right) \quad \text{(Expected enrollment)} \\
E(y|d=1) &= \int_0^{\infty} y f(y|d=1)dy \quad \text{(Conditional expected education expenditure)} \\
E(y) &= P(d=1)E(y|d=1) \quad \text{(Unconditional expected education expenditure)} \\
E(ys) &= \int_0^1 \int_0^{\infty} ysf(y,s)dyds \quad \text{(Unconditional expected core expenditure)} \\
E(ys|d=1) &= \frac{E(ys)}{P(d=1)} = \frac{E(ys)}{\Phi\left(x_d'\beta_1\right)} \quad \text{(Conditional expected core expenditure)}
\end{aligned}$$

where $f(y,s)$ is the joint probability density function for $y$ and $s$ and the subscript $i$ is omitted for simplicity. We use simulations to compute the standard errors for the equations above and evaluate only at the sample means to reduce the computational burden of numerical integrations. The details of the mathematical expressions used for numerical integrations and procedures for calculating the point estimates and standard errors of the marginal effects are described in Appendix B.

## 4 Data

We primarily use the Household Expenditure Survey (HES) for the year 1995 and Household Income Expenditure Survey (HIES) for the years 2000, 2005, and 2010, all of which are conducted by the

Bangladesh Bureau of Statistics.[5] These datasets provide detailed information on individual educational expenditure as well as socioeconomic and demographic characteristics of the households. We separately analyze each round of survey for pre-SSC secondary-school age group—which are officially ages 11 to 15.[6] In addition, HES for the year 1991 is also used for the analysis of timely graduation from secondary school but HES 1991 is not used for other analysis because it does not contain individual-level education expenditure.

There are two reasons why we primarily study on secondary education. First, as shown in Figure 2, there is a significant increment in education expenditure in grade 6 and onwards in comparison with the primary level (grades 1-5). Second, the government interventions are different for the primary and secondary levels. For example, the FSPs targeted only at girls in secondary schools, whereas the Food for Education program started in 1993 and its successor, the Primary Education Stipend program started in 2002, were open to both primary-school boys and girls. While the secondary education in Bangladesh can be divided into junior secondary (grades 6-8 or ages 11-13), secondary (grades 9-10 or ages 14-15), and higher secondary (grades 11-12 or ages 16-17), we focus on the pre-SSC level or grades 6-10 because passing the SSC examination is a major milestone in the educational attainment in Bangladesh.

The analysis of older age groups, including the higher secondary and tertiary levels, are beyond the scope of this paper, because the analysis gets more complicated for several reasons. First, early marriage and pregnancy can result in grade repetition and dropout for girls, but we have only limited information about each child beyond gender and age. As a result, our three-part model cannot address these issues and our estimates are likely to be confounded with early marriage and pregnancy. Second, the passing rate of the SSC examination was historically low, below 60% for most years before 2007 as Figure 1 shows. Since we do not have information about whether the child has passed or failed the SSC examination, our estimate of the gender effect is likely to be confounded with the results of SSC examination. Finally, the proportion of girls in higher education was very small in earlier years, making it difficult to attain reliable estimates.

We include the following set of covariates in all of our regressions for all three equations (i.e., eqs. (1)-(3)): demographic characteristics of the household such as the age and gender of the child, the age and gender of the household head, logarithmic household size, logarithmic expenditure per capita, the number of children, head's working status and religion, and parental education in years. In addition, we include urban dummy to capture the geographical heterogeneity in parental investment

[5]Top 1% observations with the highest total educational expenditure are dropped to exclude outliers.
[6]Official primary-school age group is ages 6 to 10 in Bangladesh.

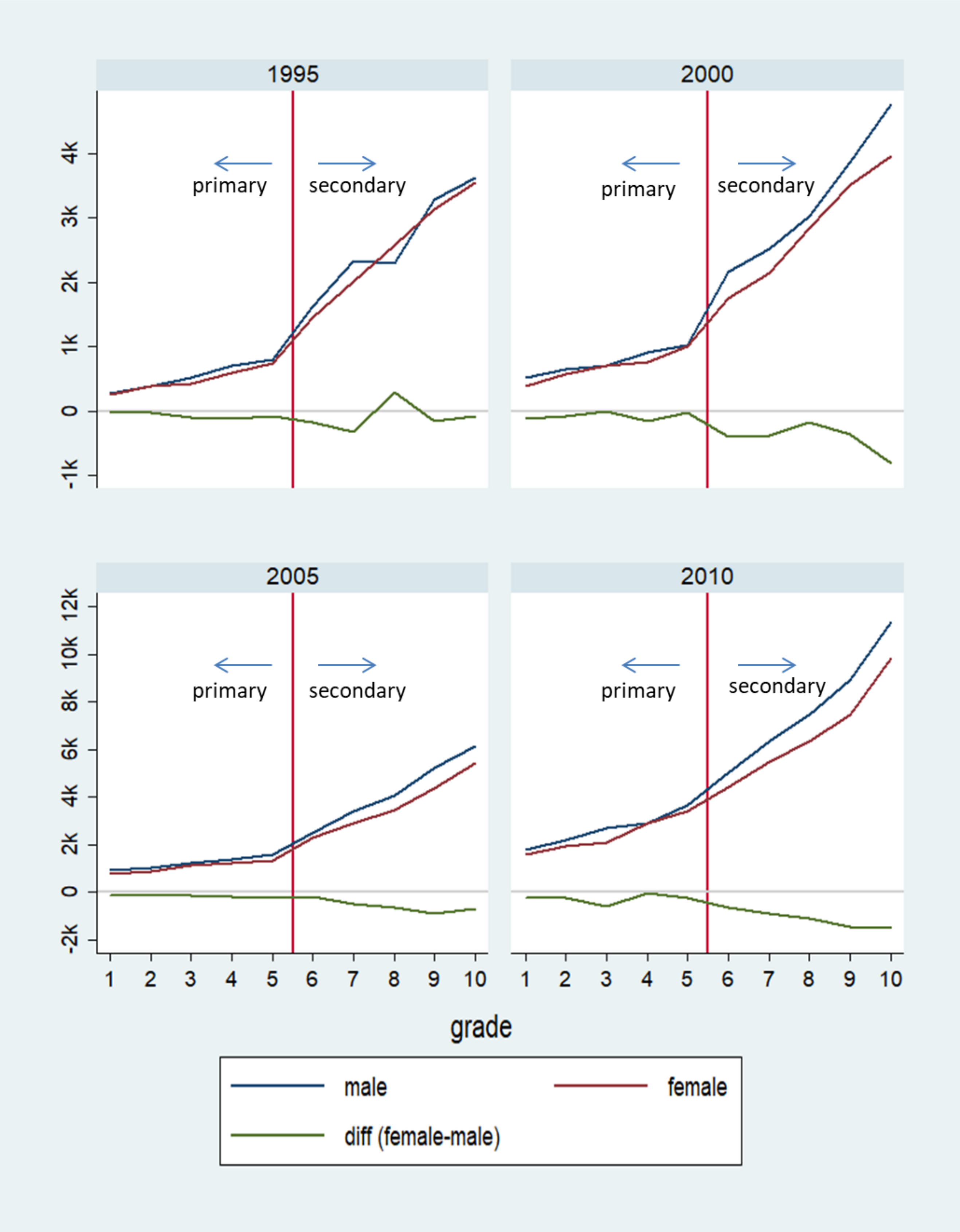


Figure 2: Nominal education expenditure in taka by year, gender, and grade

on children's education in each equation. The choice of these covariates are consistent with the existing studies such as Kingdon (2005), Aslam and Kingdon (2008), Masterson (2012), and Azam and Kingdon (2013).

We also include some variables that affect some but not all equations. Our model allows this flexibility of adopting different sets of covariates for each of eqs. (1)-(3). School accessibility may heavily affect the enrollment decision particularly in developing countries such as Bangladesh where the infrastructure and institutions are underdeveloped. Meanwhile, it is unlikely to heavily affect education expenditure. Thus, in eq. (1), we additionally include the numbers of secondary schools and madrasas per thousand people in the area of residence, which is a district for the years 1995, 2005, and 2010 and a subdivision for the year 2000, as measures of school accessibility. For the years 1995, 2005, and 2010, we obtain the numbers of secondary schools and madrasas at the district level from BANBEIS (1995), BANBEIS (2006) and BANBEIS (2010). For the year 2000, we obtain these numbers at the subdivision level from Bangladesh Bureau of Statistics (2002). We then divide the number of schools by the population figures taken from the Population and Housing Census for the year 2001. For eq. (2), school type variables [7] are added as different school types may affect tuition, uniform, and other education fees. The logarithmic education expenditure is separately added to control the education expenditure in the core share equation (i.e., eq. (3)).

Table 2 reports descriptive summary statistics for the secondary school enrollment, and basic covariates (as discussed above) disaggregated by year and gender of the child in the secondary-school age group, regardless of the enrollment status of the child for the years 1995 and 2010.[8]

The first row in Table 2 shows that girls are on average more likely to be enrolled in school. The gender difference in enrollment was small and not statistically significantly different from zero by a $t$-test of equality of means in 1995, but it has become larger and statistically significant since the year 2000. This fits well with the common observation of the reversal of the gender gap from pro-male to pro-female in school enrollment in Bangladesh in recent years (e.g., Asadullah and Chaudhury (2009)).

Table 2 also shows that girls tend to be younger and live in a larger household than boys. Over the years, there is an increase in the enrollment rate for both genders and an impressive rise in the

[7] We use three school types: government, private and all other schools. Other schools include NGO schools, madrasas and other types of schools. While the choice of school type is potentially important, we choose not to model it for three reasons. First, government schools are very rare in Bangladesh, which accounts for less than five percent of all secondary schools based on BANBEIS (1995), BANBEIS (2006) and BANBEIS (2010). Second, there is a significant mismatch in the distribution of school types between the HIES data and other sources. The proportion of children in government schools in our data is around 20 percent, which is much higher than five percent or less reported by BANBEIS or Education Watch. This discrepancy may in part stem from the public nature of private schools in Bangladesh, where private school teachers are often paid by the government. Lastly, the tuition fee reflects the quality of education as discussed below and Appendix C. It should also be noted that our results remain qualitatively similar even when the school-type variables are dropped from the regression.

[8] The summary statistics for the years 2000 and 2005 corresponding to Table 2 are reported in Table 15 in Appendix E.

Table 2: Summary statistics of basic covariates by gender for 1995 and 2010 (secondary-school age group)

| | 1995 | | | | 2010 | | | |
|---|---|---|---|---|---|---|---|---|
| | **Boy (B)** | **Girl (G)** | **Diff (G-B)** | **All** | **Boy (B)** | **Girl (G)** | **Diff (G-B)** | **All** |
| Variables | (1) | (2) | (2)-(1) | (4) | (5) | (6) | (6)-(5) | (8) |
| ***All children aged 11-15*** | | | | | | | | |
| Enrolled in secondary school | 0.349 | 0.370 | 0.021 | 0.359 | 0.465 | 0.560 | 0.095 | 0.511 |
| | (0.477) | (0.483) | | (0.480) | (0.499) | (0.496) | *** | (0.500) |
| Child's age (yrs) | 13.022 | 12.903 | -0.119 | 12.966 | 12.980 | 12.896 | -0.084 | 12.940 |
| | (1.369) | (1.351) | *** | (1.362) | (1.389) | (1.372) | ** | (1.382) |
| HH per capita expenditure | 10.222 | 11.512 | 1.29 | 10.832 | 28.434 | 28.659 | 0.225 | 28.543 |
| | (8.062) | (11.161) | *** | (9.673) | (19.044) | (21.466) | | (20.248) |
| Household size | 6.634 | 6.807 | 0.173 | 6.716 | 5.518 | 5.605 | 0.087 | 5.560 |
| | (2.507) | (2.518) | ** | (2.513) | (2.005) | (1.868) | * | (1.940) |
| Father's education (yrs) | 3.691 | 3.951 | 0.26 | 3.814 | 2.780 | 2.832 | 0.052 | 2.805 |
| | (4.426) | (4.578) | ** | (4.500) | (4.150) | (4.172) | | (4.160) |
| Mother's education (yrs) | 1.960 | 2.262 | 0.302 | 2.103 | 2.484 | 2.579 | 0.095 | 2.530 |
| | (3.085) | (3.347) | *** | (3.215) | (3.595) | (3.674) | | (3.633) |
| Number of children | 3.658 | 3.794 | 0.136 | 3.722 | 2.932 | 3.036 | 0.104 | 2.982 |
| | (1.862) | (1.913) | ** | (1.888) | (1.438) | (1.444) | *** | (1.442) |
| Urban | 0.314 | 0.365 | 0.051 | 0.338 | 0.342 | 0.335 | -0.007 | 0.339 |
| | (0.464) | (0.482) | *** | (0.473) | (0.474) | (0.472) | | (0.473) |
| Female head | 0.084 | 0.089 | 0.005 | 0.086 | 0.131 | 0.139 | 0.008 | 0.135 |
| | (0.277) | (0.285) | | (0.281) | (0.337) | (0.346) | | (0.342) |
| Head is a wage worker | 0.354 | 0.365 | 0.011 | 0.359 | 0.407 | 0.404 | -0.003 | 0.405 |
| | (0.478) | (0.482) | | (0.480) | (0.491) | (0.491) | | (0.491) |
| Head's age (yrs) | 46.466 | 46.556 | 0.09 | 46.508 | 47.142 | 46.827 | -0.315 | 46.990 |
| | (11.188) | (11.115) | | (11.152) | (10.597) | (10.554) | | (10.577) |
| Muslim | 0.898 | 0.890 | -0.008 | 0.894 | 0.898 | 0.887 | -0.011 | 0.892 |
| | (0.303) | (0.313) | | (0.308) | (0.303) | (0.317) | | (0.310) |
| Hindu | 0.094 | 0.101 | 0.007 | 0.097 | 0.093 | 0.103 | 0.010 | 0.098 |
| | (0.292) | (0.301) | | (0.296) | (0.290) | (0.305) | | (0.297) |
| *Obs* | 2,641 | 2,370 | | 5,011 | 3,209 | 2,996 | | 6,205 |
| ***Enrolled in secondary school children aged 11-15*** | | | | | | | | |
| Govt school | 0.16 | 0.18 | 0.02 | 0.17 | 0.23 | 0.20 | -0.03 | 0.22 |
| | (0.37) | (0.39) | | (0.38) | (0.42) | (0.40) | * | (0.41) |
| Private school | 0.79 | 0.81 | 0.02 | 0.80 | 0.70 | 0.69 | -0.01 | 0.70 |
| | (0.41) | (0.40) | | (0.40) | (0.46) | (0.46) | | (0.46) |
| Other school | 0.05 | 0.01 | -0.04 | 0.03 | 0.07 | 0.10 | 0.03 | 0.09 |
| | (0.21) | (0.11) | *** | (0.17) | (0.26) | (0.30) | *** | (0.28) |
| *Obs* | 921 | 877 | | 1,798 | 1,493 | 1,679 | | 3,172 |

Note: Standard errors are reported in parentheses below the mean. ***,** ,* denote that the means of girl and boy are different at 1, 5, 10 percent significance levels, respectively. The unit for household per capita expenditure is thousand taka. Other school includes all types of schools other than government and private schools, including religious schools (like madrasas) and NGO schools.

nominal household per capita expenditure, where it has almost tripled in 2010 compared with that in 1995. There are improvements shown by other indicators such as an increase in mother's education.

Next, we analyze the pattern of education expenditure using the subsample of children who were enrolled in secondary school at the time of survey. Table 3 reports summary statistics of each education expenditure item in nominal terms for the years 1995 and 2010.[9] The table shows that the average education expenditure has rapidly increased.[10]

Figure 2 further demonstrates the trend in gender disparity of average education expenditure based on the grade of child enrolled in school at the time of survey for each year. There are three points to note from this figure. First, boys are getting a larger share of educational investment than girls conditional on enrollment. Second, the gap is widening for higher grades except 1995, especially at the secondary level (grade 6 onwards). Finally, the education expenditure has a positive correlation with grade and increased over years for both genders.

As mentioned earlier, we categorize the items of education expenditure into core and peripheral components. The core component includes tuition, home tutoring, and materials, where materials refer to expenses on textbooks, exercise books, and stationary. The peripheral component includes the other recorded items: admission, examination, uniform, meals, transportation, and others, which are typically perceived to have only a marginal relevance at best to the quality of education.

It is reasonable to include the tuition fee in the core component because it appears to reflect, at least to some extent, the quality of education provided by the schools in Bangladesh. If schools face some degree of competition, those schools which consistently provide only low-quality education at a high tuition will exit the market such that a positive correlation between the quality of education and tuition would emerge. As elaborated in Appendix C, an analysis of separate dataset provides suggestive evidence that a higher tuition reflects higher quality of education based on the relationship between the average tuition fee and test score at the primary level.

We also include home tutoring in the core expenditure. It is widely documented that home tutoring can be an important educational input and this is also the case in Bangladesh. It is not uncommon in Bangladesh for public school teachers to serve as private tutors for their students. In some cases, the teachers deliberately teach less in the regular classes to gain more incomes from private tutoring. Given such a possibility, home tutoring must be included into the core component.

Nevertheless, a concern about the interpretation of the spending on home tutoring may arise here.

[9] The same summary statistics for the years 2000 and 2005 are reported in Table 16 in Appendix E.

[10] The rate of increase for education expenditure has been faster than the inflation rate. For example, between 2005 and 2010, the average annual inflation rate in consumer prices was 8.6 percent based on World Bank statistics, and the annual rate of increase in secondary education expenditure was around 17 percent in the same period.

On one hand, home tutoring would raise the overall education quality that the child receives. On the other hand, if home tutoring is given only to weaker students and boys are generally weaker than girls, the pro-male bias in the core share we show subsequently may be driven by the relatively weak academic performance of boys. However, we argue that this is highly unlikely given that girls have underperformed boys in the passing rate and share of top students in the SSC exam over years as shown in Figure 1 earlier.

Finally, it is also reasonable to include materials in the core component, because reading more textbooks and doing more exercises also directly contribute to the academic performance. However, one could argue that more expensive books are not necessarily of higher quality. Thus, the inclusion of materials in the core component is admittedly disputable. Therefore, we repeated our analysis excluding the materials from the core component (unreported) but the results are qualitatively similar. In sum, our choice of the definition of the core component is reasonable, if not undisputable, and consistent with available empirical and anecdotal evidence.

Table 3: Summary statistics of education expenditure by items for secondary-school enrollees in 1995 and 2010

| | **1995** | | | | **2010** | | | |
|---|---|---|---|---|---|---|---|---|
| Taka | **Boy (B)** (1) | **Girl (G)** (2) | **Diff (G-B)** (2)-(1) | **% Zeros** (4) | **Boy (B)** (5) | **Girl (G)** (6) | **Diff (G-B)** (6)-(5) | **% Zeros** (8) |
| **Core** | 1,672.7 | 1,582.1 | -90.6 | 1% | 5,239.4 | 4,284.8 | -954.6 | 0% |
| | (1,616.3) | (1,539.7) | | | (5,081.9) | (4,362.9) | *** | |
| *Tuition* | 275.0 | 193.7 | -81.3 | 32% | 548.7 | 296.2 | -252.5 | 46% |
| | (312.7) | (304.9) | *** | | (963.1) | (606.0) | *** | |
| *Home Tutor* | 802.8 | 788.8 | -14 | 45% | 3,273.2 | 2,626.5 | -646.7 | 26% |
| | (1,298.1) | (1,225.6) | | | (4,334.5) | (3,745.4) | *** | |
| *Material* | 594.9 | 599.5 | 4.6 | 1% | 1,417.5 | 1,362.1 | -55.4 | 1% |
| | (429.2) | (414.8) | | | (950.1) | (929.4) | * | |
| **Peripheral** | 717.2 | 747.3 | 30.1 | 1% | 2,109.8 | 2,066.8 | -43 | 0% |
| | (877.9) | (791.3) | | | (2,223.5) | (2,076.6) | | |
| *Admission* | 126.4 | 138.4 | 12 | 24% | 371.2 | 336.6 | -34.6 | 21% |
| | (211.3) | (196.8) | | | (657.3) | (561.2) | | |
| *Exam* | 115.2 | 123.7 | 8.5 | 5% | 301.3 | 295.1 | -6.2 | 5% |
| | (145.8) | (138.9) | | | (287.9) | (270.3) | | |
| *Uniform* | 215.2 | 249.3 | 34.1 | 45% | 618.8 | 629.5 | 10.7 | 19% |
| | (289.6) | (278.4) | ** | | (534.4) | (657.7) | | |
| *Meal* | 40.3 | 4.9 | -35.4 | 99% | 423.9 | 377.3 | -46.6 | 58% |
| | (463.6) | (57.7) | ** | | (805.8) | (744.1) | * | |
| *Transportation* | 87.0 | 109.2 | 22.2 | 81% | 204.7 | 311.4 | 106.7 | 85% |
| | (332.7) | (393.8) | | | (817.7) | (1,079.6) | *** | |
| *Others* | 133.0 | 121.7 | -11.3 | 44% | 190.0 | 116.9 | -73.1 | 75% |
| | (281.3) | (343.9) | | | (1,272.9) | (775.7) | * | |
| **Total** | 2,389.9 | 2,329.4 | -60.5 | | 7,349.2 | 6,351.6 | -997.6 | |
| | (2,111.5) | (2,030.0) | | | (6,150.7) | (5,524.3) | *** | |
| **Core Share** | 0.68 | 0.65 | -0.03 | | 0.67 | 0.63 | -0.04 | |
| | (0.19) | (0.20) | *** | | (0.18) | (0.19) | *** | |
| Obs | 921 | 877 | | | 1,493 | 1,679 | | |

Note: Standard errors are reported in parentheses below the mean. ***,**,* denote that the means of girl and boy are different at 1, 5, 10 percent significant levels, respectively. The summary statistics is for subsample of children who were enrolled in school at the time of survey. Core share stands for the ratio of core components over total education expenditure. The annual session and registration fees are also included in admission because they are not separately reported in HES 1995.

Table 3 presents the descriptive statistics for each education expenditure item for the years 1995 and 2010. From the table we see that the core component accounts roughly two-thirds of the total education expenditure and boys have a significantly higher share than girls. Within the core component, home tutoring fee is the major cost item but a considerable share of children have no spending on home tutor in both years. There is an obvious trend in the popularity of home tutoring over the years, particularly among higher grades. In 1995, 45% of secondary students reported to have no private tutors, but this ratio dropped to 26% in 2010, showing increasing dependency on home tutors. This may also indicate that parents are willing to invest more in children's education for better quality of education beyond the typical costs like schooling fees.[11] Table 3 also shows that girls on average have lower spending on tuition and a significant share of children have zero spending on tuition fee (32% in 1995 and 46% in 2010), which can be explained by the tuition waiver provided by various programs including the FSPs discussed in detail in Section 6.

# 5 Main Results

In this section, we present our main results. We first show the ML estimates of the three-part model. We then perform similar regressions under alternative specifications to show the robustness of our results. Finally, we compute the marginal effects of being a girl with the method discussed at the end of Section 3 to provide results with direct quantitative interpretations.

### Estimation of coefficients

Table 4 presents the ML estimates of the coefficient on the girl dummy in the three-part model for each year and for each of primary- and secondary-school age groups.[12] It is shown that the gender gap for the primary age group is smaller than that for the secondary-school age group, and thus we hereafter focus on the analysis of the secondary-school age group.

Columns (4) to (6) of Table 4 show the presence of clear and strong pro-female bias in enrollment decision from the year 2000 onwards after controlling for the observables discussed in Section 4. That is, other things being equal, parents are more likely to send girls to school than boys. In contrast, conditional on enrollment, the core component for girls tends to account for a lower share of the total education expenditure than that for boys. Column (5) reveals that, conditional on enrollment,

[11] Of course, alternative interpretations are possible. For example, the increasing popularity of home tutoring may reflect the deteriorating quality in school education because of overcrowding of classrooms or teacher absenteeism (Banerjee and Duflo, 2006).

[12] Equation by equation regressions (i.e., under the assumption of uncorrelated errors where all the $\rho$'s are zero) yield similar results and are presented in Table 18 in Appendix E. The importance of introducing the dependent error structure is also discussed in Appendix E.

households are spending significantly less on girls' education than boys' for all the four years. The gender gap in 1995 somewhat differs from the three more recent rounds. While we still see pro-male bias in conditional education expenditure, the coefficient on the girl dummy is substantially smaller in absolute value and the coefficients for enrollment and core share equations are insignificant. We will revisit this issue in Section 7.

While the girl dummy is the main covariate of interest, other covariates included in our regressions are also of interest. Therefore, we briefly summarize our findings here. The details of the regressions presented in Table 4 are reported in Table 17 in Appendix E. In general, children in richer households are more likely to be enrolled and receive a higher expenditure on education but a lower core share. Parental education, especially mother's education, has a similar effect qualitatively in all three decisions. The more educated parents are, the more likely children will enroll in school and get more education expenditure, despite a lower core share, indicating the presence of positive intergenerational transmission in education. Somewhat surprisingly, the number of children has no effect as most of its coefficients are not significant. The difference between urban and rural areas exhibits an interesting pattern. Children in rural areas are more likely to enroll in school, but have lower education expenditure conditional on enrollment. These differences may be caused by various aid programs targeted to rural areas. If the head is a wage worker, child has a lower probability of attending school. Other covariates, such as the logarithm of household size, head's sex, age, and religion, are not statistically significant. The coefficients on school-type variables show that children going to private schools spend more on education than those going to government schools. As to the core share decision, the estimated coefficients on the logarithmic education expenditure are all positive but insignificant. Finally, the importance of school accessibility in affecting enrollment decision is worth highlighting. When the number of secondary schools per thousand people in the area of residence is higher, children are more likely to enroll, while the accessibility for madrasas does not seem to have much impact.

Because of the grade repetition and delayed entry into school, some secondary-school age children may be still in primary school and some post-secondary-school-age children may be still in secondary school. To see if the presence of these children affects our results, we re-estimate the same model with an alternative definition of age groups where primary- and secondary-school age groups are defined as 6-11 and 12-17, respectively. The results are quantitatively and qualitatively similar.[13]

To understand the time trend of the gender bias in education expenditure, we estimated the three-part model for all years simultaneously with time fixed effect and its interaction term with the girl dummy. As the regression results with the pooled sample in Table 5 show, the gender bias pattern

[13] Results are available upon request.

Table 4: ML estimation of the three-part model by years and age groups

| | **Primary-school age (6-10)** | | | **Secondary-school age (11-15)** | | |
|---|---|---|---|---|---|---|
| | $d$ | Cond $y$ | Cond $s$ | $d$ | Cond $y$ | Cond $s$ |
| *Coef.* | (1) | (2) | (3) | (4) | (5) | (6) |
| **1995** | | | | | | |
| Girl | -0.031 | -0.013 | -0.016 | -0.001 | -0.085*** | 0.001 |
| | (0.036) | (0.033) | (0.012) | (0.042) | (0.032) | (0.032) |
| *Obs.* | | 6485 | | | 5011 | |
| **2000** | | | | | | |
| Girl | 0.061* | -0.114*** | 0.009 | 0.339*** | -0.174*** | -0.082*** |
| | (0.036) | (0.036) | (0.010) | (0.039) | (0.049) | (0.014) |
| *Obs.* | | 5600 | | | 4878 | |
| **2005** | | | | | | |
| Girl | 0.048 | -0.076** | -0.023** | 0.291*** | -0.154*** | -0.071*** |
| | (0.035) | (0.033) | (0.009) | (0.034) | (0.027) | (0.012) |
| *Obs.* | | 6481 | | | 5638 | |
| **2010** | | | | | | |
| Girl | 0.134*** | -0.066** | -0.019* | 0.289*** | -0.131*** | -0.067*** |
| | (0.032) | (0.029) | (0.010) | (0.033) | (0.025) | (0.009) |
| *Obs.* | | 7272 | | | 6205 | |

Note: ***,**,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at the household level are reported in parentheses. The estimations are obtained using three-part model constructed in Section 3. In all regressions, the following covariates are also included: logarithmic per capita expenditure, logarithmic household size, father's and mother's education in years, number of children, female head, wage-worker head, head's age, and religion (muslim/hindu), and urban area. In addition, secondary school and madrasa school accessibility variables, school types (government/private school) and logarithmic education expenditure are also controlled in $d$, Cond $y$ and Cond $s$, respectively. Detailed results for secondary-school age group are presented in Table 17 in Appendix E.

detected are consistent with the year-by-year results. That is, a pro-female bias is found in enrollment decision but a pro-male bias is found in conditional expenditure and core share decisions. The coefficients on the interaction terms between the year and girl dummy show that enrollment decision has become more pro-female. On the contrary, the core share has become more pro-male. The conditional expenditure did not change much over time.

Table 5: Results of the pooled regression with the three-part model

| *Coef.* | $d$ (1) | Cond $y$ (2) | Cond $s$ (3) |
|---|---|---|---|
| Girl | 0.029 | -0.097*** | -0.032*** |
| | (0.040) | (0.033) | (0.010) |
| $Y_{00}$ | -0.036 | 0.224*** | -0.017 |
| | (0.039) | (0.035) | (0.011) |
| $Y_{05}$ | -0.042 | 0.400*** | -0.037*** |
| | (0.040) | (0.032) | (0.013) |
| $Y_{10}$ | -0.161*** | 0.541*** | -0.054*** |
| | (0.045) | (0.035) | (0.016) |
| Girl $\times Y_{00}$ | 0.317*** | -0.059 | -0.050*** |
| | (0.055) | (0.047) | (0.015) |
| Girl $\times Y_{05}$ | 0.259*** | -0.072* | -0.034** |
| | (0.053) | (0.042) | (0.014) |
| Girl $\times Y_{10}$ | 0.260*** | -0.038 | -0.032** |
| | (0.052) | (0.041) | (0.013) |
| Obs | | 21,732 | |

Note: ***,**,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at the household level are reported in parentheses. Additional controls include the set of covariates discussed in Table 4 except that the school accessibility variables are constructed at subdivision level for all years to make this variable comparable across years. Year 1995 is the base year for comparison in these regressions.

Therefore, Table 5 indicates that the apparent inconsistency in the direction of gender bias did not change and, if any thing, strengthened by the fact that the pro-female bias in enrollment became stronger while the pro-male bias stayed the same in conditional expenditure and were strengthened in core share decision.

### Robustness of Estimation

There are some endogeneity concerns, which may bias our estimation, in the results presented above. Therefore, we perform similar regressions under alternative specifications that would mitigate these concerns to show that our estimates are robust to these endogeneity concerns. We also conduct the analysis separately for urban and rural areas to show that our main findings hold in both areas.

To understand the endogeneity concern, recall that Table 2 shows that girls on average live in a household with significantly more children and household members than boys. This may be explained by the fertility stopping rule with unobserved parental preference towards boys (Jensen, 2002). That is, if parents have a preference for a boy, they may continue to try to have more children until they have a boy, leading to a bigger family size for girls on average. Hence, the unobserved parental preference may simultaneously affect both the household's demographic composition as well as the education expenditure on children such that the unobserved error terms may be correlated with the covariates.

To partially address this concern, we include the household size and number of children in the set of covariates to control for the differences in the household structure in our regressions. However, these controls would not fully address the potential endogeneity concerns relating to the family structure. Therefore, we run linear regressions with household fixed effects to control for all household-level observable and unobservable characteristics in addition to the individual-level observable characteristics. The signs of the coefficient on the girl dummy variable from these estimations are broadly consistent as can be seen from Table 6, though the level of significance drop for the conditional expenditure decision. This may be partly because of the smaller sample size as we use subsample of children from households with at least two children in secondary-school age group.

Another concern regarding this family structure issue is that girls are likely to face a stiffer competition with siblings than boys because the former have more siblings than the latter on average. Therefore, our main results may be driven by the differential competitions for boys and girls. To mitigate this issue, we also analyze a subsample of households in which there is only one child. This arguably mitigates the gender difference in the level of competition within the household. Because of the small sample size used for this analysis, it is difficult to draw definitive conclusions, but the results reported in Columns (1) to (3) of Table 7 indicate that the biases remain but become weaker. Similar results are obtained when we restrict the sample to be children living in households with one boy and one girl in secondary-school age group. Therefore, the competition within the household appears to be a part of the source of bias, though the explanatory power is limited.

Because of the differences between urban and rural in economic environment, labor market development, and social attitudes towards female education among others, one may argue that rural and urban areas should be separately analyzed. Moreover, as shown in Section 6, the FSPs only covered non-metropolitan areas. Thus, we re-estimate the analysis of the three-part model separately for the

Table 6: Results of linear regressions with household-level fixed effects

| *Coef.* | $d$ (1) | Cond $y$ (2) | Cond $s$ (3) |
|---|---|---|---|
| **1995** | | | |
| Girl | -0.006 | -0.139* | -0.014 |
| | (0.028) | (0.076) | (0.027) |
| *Obs* | 2,834 | 1,076 | 1,076 |
| **2000** | | | |
| Girl | 0.076*** | -0.063 | -0.043* |
| | (0.028) | (0.090) | (0.025) |
| *Obs* | 2,695 | 1,015 | 1,015 |
| **2005** | | | |
| Girl | 0.098*** | -0.032 | -0.018 |
| | (0.028) | (0.068) | (0.015) |
| *Obs* | 2,587 | 1,084 | 1,084 |
| **2010** | | | |
| Girl | 0.095*** | -0.078 | -0.050*** |
| | (0.031) | (0.061) | (0.019) |
| *Obs* | 2,551 | 1,220 | 1,220 |

Note: ***,** ,* denote statistical significance at 1, 5, 10 percent levels. Standard errors are reported in parentheses. Each point estimate corresponds to one linear regression. Household-level fixed-effects terms as well as the age fixed effects are included in all regressions. In addition, school type dummies for government and private school are controlled in column (2) and logarithmic education expenditure is added in column (3). All other covariates are absorbed in the household-level fixed effects.

Table 7: Linear regressions by subsamples with different family structure

| | **Only Child** | | | **One-boy-one-girl** | | |
|---|---|---|---|---|---|---|
| | $d$ | Cond $y$ | Cond $s$ | $d$ | Cond $y$ | Cond $s$ |
| *Coef.* | (1) | (2) | (3) | (4) | (5) | (6) |
| **1995** | | | | | | |
| Girl | 0.023 | 0.097 | -0.056* | 0.010 | -0.139 | 0.001 |
| | (0.052) | (0.151) | (0.032) | (0.033) | (0.096) | (0.041) |
| *Obs* | 314 | 113 | 113 | 1,076 | 423 | 423 |
| **2000** | | | | | | |
| Girl | 0.064 | -0.130 | -0.013 | 0.069** | -0.135 | -0.044 |
| | (0.052) | (0.142) | (0.038) | (0.032) | (0.091) | (0.029) |
| *Obs* | 286 | 108 | 108 | 1,146 | 447 | 447 |
| **2005** | | | | | | |
| Girl | 0.025 | -0.129 | -0.042 | 0.099*** | -0.037 | -0.022 |
| | (0.048) | (0.095) | (0.028) | (0.032) | (0.077) | (0.017) |
| *Obs* | 382 | 169 | 169 | 1,190 | 526 | 526 |
| **2010** | | | | | | |
| Girl | 0.040 | -0.089 | -0.046** | 0.093*** | -0.068 | -0.054** |
| | (0.038) | (0.076) | (0.018) | (0.035) | (0.070) | (0.022) |
| *Obs* | 580 | 305 | 305 | 1,086 | 510 | 510 |
| *Basic covariates* | Y | Y | Y | Y* | Y* | Y* |
| *HH fixed effects* | N | N | N | Y | Y | Y |

Note: ***,** ,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at household levels are reported in parentheses. The estimations are obtained by equation-by-equation OLS estimations for each dependent variable. Only child subsample contains child from the household with only one child. One-boy-one-girl subsample contains children from the households with exactly two children, one secondary-school age boy and one secondary-school age girl.
*: The girl dummy and age fixed effects are included in columns (4)-(6). In addition, the school type dummies and logarithmic education expenditure are included, respectively, in columns (5) and (6). All other covariates are absorbed in the household-level fixed effects.

Table 8: Estimation of the three-part model by the urban and rural subsamples

| | Urban | | | Rural | | |
|---|---|---|---|---|---|---|
| *Coef.* | *d* | Cond *y* | Cond *s* | *d* | Cond *y* | Cond *s* |
| **1995** | | | | | | |
| Girl | 0.094 | 0.008 | -0.030* | -0.047 | -0.131*** | 0.010 |
| | (0.072) | (0.047) | (0.016) | (0.053) | (0.043) | (0.047) |
| *Obs* | | 1,695 | | | 3,316 | |
| **2000** | | | | | | |
| Girl | 0.310*** | -0.024 | -0.047** | 0.365*** | -0.277*** | -0.116*** |
| | (0.073) | (0.053) | (0.021) | (0.047) | (0.059) | (0.019) |
| *Obs* | | 1,598 | | | 3,280 | |
| **2005** | | | | | | |
| Girl | 0.264*** | -0.102** | -0.054*** | 0.318*** | -0.177*** | -0.081*** |
| | (0.060) | (0.046) | (0.016) | (0.042) | (0.034) | (0.016) |
| *Obs* | | 1,921 | | | 3,717 | |
| **2010** | | | | | | |
| Girl | 0.376*** | -0.095** | -0.069*** | 0.255*** | -0.151*** | -0.069*** |
| | (0.057) | (0.043) | (0.015) | (0.041) | (0.032) | (0.011) |
| *Obs* | | 2,102 | | | 4,103 | |

Note: ***,** ,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at the household level are reported in parentheses. The same set of covariates is used as in Table 4 except that the urban dummy is dropped.

urban and rural areas. As the results in Table 8 show, the directions of the gender gap in three equations are essentially the same except that they are less pronounced in 1995. The comparison of the two areas also shows that the gender gap in rural areas is clearer than that in urban area.

### Marginal Effects

Because the regression coefficients do not translate into a readily interpretable quantity, we evaluate the marginal effect of being a girl at the sample mean using the formulae presented at the end of Section 3. The estimated marginal effects are presented in Table 9. Column (1) shows the difference between girls and boys in probability of enrollment. There is a significant pro-female bias except in 1995. In 2010, girls are 11.6 percentage points more likely to enroll in secondary schools than boys at the sample mean. The effects on conditional expectations are shown in columns (3) and (5) and male students apparently have an advantage over female students in having more education expenditure and higher amount spent on the core components. For example, the differences in the total education expenditure between boys and girls in 2005 was 416.6 taka at the mean of the subsample of secondary-school enrollees. Similarly, there exists a significant pro-male bias in the core component expenditure from 2000 onwards. However,

Table 9: Marginal effects of the girl dummy at the sample mean

| *Marginal effects on* | E($d$) | E($y$) | E($y \mid d = 1$) | E($ys$) | E($ys \mid d = 1$) |
|---|---|---|---|---|---|
| Year | (1) | (2) | (3) | (4) | (5) |
| **1995** | -0.001 | -40.5 | -181.9*** | -7.8 | -110.5 |
| | ( 0.016) | ( 26.3) | ( 67.7) | ( 16.3) | ( 92.7) |
| *Obs.* | 5011 | 5011 | 1798 | 5011 | 1798 |
| **2000** | 0.126*** | 152.5*** | -224.7*** | 11.5 | -312.7*** |
| | ( 0.014) | ( 29.8) | ( 76.3) | ( 24.6) | ( 62.5) |
| *Obs.* | 4878 | 4878 | 1885 | 4878 | 1885 |
| **2005** | 0.114*** | 145.6*** | -416.6*** | -0.4 | -367.3*** |
| | ( 0.014) | ( 47.6) | ( 80.8) | ( 40.3) | ( 56.6) |
| *Obs.* | 5638 | 5638 | 2579 | 5638 | 2579 |
| **2010** | 0.116*** | 313.0*** | -616.8*** | 3.2 | -604.9*** |
| | ( 0.014) | ( 80.6) | (146.7) | ( 51.7) | ( 98.7) |
| *Obs.* | 6205 | 6205 | 3172 | 6205 | 3172 |

Note: ***,** ,* denote statistical significance at 1, 5, 10 percent levels. Standard errors obtained by simulation with 100 replications are reported in parentheses. E(·) stands for the expectation of the variable in the brackets. Estimates in column (1) are the marginal effect of the girl dummy on the expected enrollment in secondary school for the children in the secondary-school age group. The marginal effects presented in Columns (2) to (5) are in taka in nominal terms. Unconditional [Conditional] expectations are evaluated at the mean of the full sample [subsample of secondary-school enrollees].

as shown in column (2), when we combine the probability of enrollment and conditional expenditure effect together, girls have a higher unconditional education expenditure than boys on average except 1995. Further, the gender gap in the unconditional core education expenditure became negligible as shown in column (4). This highlights the importance of decomposing the education expenditure decision into different parts.

The results above consistently show that girls received less expenditure in the core component than boys conditional on enrollment across years, and this gender gap grew over time. However, because the core component consists of multiple items, including tuition, private tutoring, and materials, we further investigate this gender gap by item-by-item Tobit regressions. We report the marginal effect of being a girl at the sample mean for the secondary-school enrollees in Table 10. We find that girls receive significantly less investment in tuition than boys for all the survey years. Girls also receive less in home tutoring, though not all the differences are significant. On the other hand, the only item for which girls

Table 10: Tobit marginal effect of the girl dummy on education expenditure by expenditure item among secondary-school enrollees

| Expenditure in Taka | **1995** | **2000** | **2005** | **2010** |
|---|---|---|---|---|
| **Core** | -178.7*** | -284.1*** | -259.8*** | -649.9*** |
| | (62.7) | (70.9) | (77.4) | (137.6) |
| *Tuition* | -228.9*** | -488.0*** | -694.6*** | -669.0*** |
| | (26.6) | (38.4) | (60.8) | (63.5) |
| *Home Tutor* | -142.7 | -199.1* | -100.1 | -578.8*** |
| | (87.4) | (101.9) | (108.2) | (153.6) |
| *Material* | 1.7 | -5.4 | -23.1 | -14.9 |
| | (19.3) | (21.1) | (20.5) | (31.1) |
| **Peripheral** | 6.4 | 31.0 | -45.0 | 59.8 |
| | (35.1) | (37.5) | (45.5) | (69.6) |
| *Admission* | 8.8 | -20.5 | -15.0 | -26.9 |
| | (11.5) | (13.0) | (15.5) | (24.8) |
| *Exam* | 6.9 | -2.3 | 9.6 | -1.0 |
| | (6.4) | (6.7) | (6.2) | (10.2) |
| *Uniform* | 70.0*** | 86.5*** | 25.3 | 49.1* |
| | (22.7) | (22.5) | (23.8) | (25.9) |
| *Meal* | -310.6 | 44.9 | -52.4 | -59.5 |
| | (840.1) | (37.4) | (40.7) | (57.7) |
| *Transport* | 9.2 | -7.8 | 57.7 | 723.8*** |
| | (65.8) | (95.3) | (109.7) | (187.7) |
| **Obs** | 1,798 | 1,885 | 2,579 | 3,172 |

Note: ***,** ,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at the household level are reported in parentheses. Marginal effects using Tobit regressions of education expenditure items evaluated at the mean of the subsample of secondary-school enrollees are reported. The covariates are the same to those used in columns (2) and (5) of Table 4. The annual session and registration fees are also included in admission because they are not separately reported in HES 1995.

somewhat consistently receive a higher amount is uniform but this difference does not make up for the disadvantages in other expenditure items. Therefore, girls have overall lower education expenditure and lower core expenditure conditional on enrollment and this female disadvantage mainly comes from tuition and home tutoring.

# 6 Analyzing the Role of FSPs

Given the apparent inconsistency in the direction of gender bias we found in the previous section, a natural question that arises here is what explains this inconsistency. Because a nationwide female-targeted stipend program was implemented in Bangladesh during our study period and because clear pro-male bias was observed in India and Pakistan, where no such program was implemented, we conjecture that the FSPs may have played some role and explore this possibility in this section.

First, we provide a brief background of the FSPs. We then study the impact of FSPs in four aspects. First, we incorporate the FSPs information in the three-part model constructed in Section 3 by exploiting the individual status of the receipt of FSPs and the regional variations in the FSP intensity. We show that the FSP recipients receive more educational investment conditional on enrollment than nonrecepients, but have a smaller share in the core component. By further including the girl recipient ratio (GRR), or the number of FSPs recipients over the total number of girls at the same age and in the same division of residence, and its interaction with the girl dummy, we show that girls are more likely to enroll in areas with a higher FSP intensity but receive lower educational investment in the core component than boys. Second, by exploiting the fact that the FSPs was rolled out nationwide only in 1994 and only covers for girls' secondary education, we use a triple difference strategy to identify the impact of the FSPs on school enrollment. Third, we hypothetically mute the effect of tuition waiver for girls, which is one of the major instruments of the FSPs and show that tuition waiver is an important policy instrument by which the parents are incentivized to send their girls to secondary schools. All these analyses indicate the relevance of the FSPs in explaining the gender bias we have detected in education expenditure. Finally, we explore the impact of FSPs on gender gap in education using timely graduation from secondary school as an indicator of educational outcome. The results suggest that the FSPs did not fill the gender gap in this outcome among the primary-school graduates.

### Background of FSPs

The FSPs rolled out nationwide in 1994 consist of four projects, including the Female Secondary School Assistance Project funded by International Development Association and GOB, Female Secondary Stipend Project funded by GOB, Secondary Education Development Project funded by Asian Development Bank and GOB, and Female Secondary Education Project funded by Norwegian Agency for Development Cooperation (NORAD).

With the goal of encouraging female education, all FSPs provide financial assistance to eligible female students in grades 6 to 10. They only cover unmarried girls studying in secondary schools which have signed a Cooperation Agreement with the GOB, and metropolitan areas are excluded. At the entry points (grades 6 and 9), all female students from eligible institutions are eligible to benefit from the FSPs regardless of her past attendance or performance. However, three conditions must be met for the continuation of the program; the recipients are required to i) attend at least 75 percent of the school days, ii) secure a minimum of 45 percent marks in the annual school exam, and iii) remain unmarried until the SSC examination.

The FSPs disburse the financial assistance through commercial banks in two parallel installments per academic year.[14] Based on Bangladesh Ministry of Education (1996), the stipend money, together with the book allowance if the recipient is in grade 9 and the examination fee if she is in grade 10, is disbursed to the recipients' personal bank account. When the money is withdrawn from the bank, the girls have to go personally to the assigned venue on a fixed date. The FSP recipients are also entitled to enjoy free tuition and schools are directly paid by the FSPs. However, around 15 percent of the FSP recipients pay some tuition fee in our data, even though the amount paid by them is lower than nonrecepients.

It should also be highlighted that students in grade 8 were not covered in the FSPs in 1995, because the FSPs were not rolled out at the time when they are at the entry points. Further, the FSP coverage in 1994 was substantially lower. As a result, the coverage of students in other grades also appear to be lower than later years. According to BANBEIS (2006), the number of FSP recipients was only 70 thousand in 1994. The number jumped to 1.4 million in 1995 and more than doubled in the following two years. It continued to increase rapidly until reaching its peak of 4.2 million in 2002 after which it dropped to 2.3 million in 2005. These numbers are significant both in absolute terms and relative to the size of cohort (17.3 million in 2005) and enrollment (7.4 million in 2005) for the secondary-school age group.

With the intention of improving the quality of education and reaching out to the poor regardless of gender, the FSPs were subsequently replaced by the Secondary Education Quality and Access Enhancement Program (SEQAEP) in December, 2007. Thus, FSPs are relevant only to the first three rounds of our data, namely 1995, 2000, and 2005, whereas the SEQAEP was in place in 2010. As with the FSPs, the policy on the paper for the SEQAEP does not appear to have been strictly adhered to. For example, the quota for female recipients in the SEQAEP was supposed to be 60 percent but girls account for 80 percent of the recipients in our data.

Because of the lack of clarity in the way the resources for the FSPs and SEQAEP are allocated and because of the lack of information on the individual eligibility of these programs in our dataset, our analysis is necessarily based on the actual receipt of the program. Furthermore, clean identification of the impacts of FSPs (and SEQAEP) is difficult for two reasons. First, the assignment of FSPs is

[14]The monthly stipend amount increases with grade progression, which starts from 25 taka for grade 6 and reaches 60 taka for grade 10. It roughly covered less than half of expenditure for secondary education. The tuition rate also increases with grade, from 10 taka to 15 taka for government schools grade 6 to 10, respectively. The tuition rate is 5 taka per month more for private schools. By design, it was supposed to cover tuition cost for recipients. The detailed stipend and tuition rates are displayed in Table 2 of Bangladesh Ministry of Education (1996).

nonrandom as there are some eligibility criteria. Second, we have limited data for the pre-FSPs period. In particular, individual-level information on education expenditure is only available from the year 1995 when the FSPs already started. Nevertheless, we provide some suggestive evidence that FSPs are indeed a factor contributing to the inconsistency in the direction of gender bias in Bangladesh.

### Incorporating FSPs information

To try to understand the impact of FSP at the individual level, we use the HIES data for the years 2000 and 2005 as they contain information on the individual status of the receipt of FSPs.[15]

The educational expenditure on the FSPs recipients are affected because they enjoy tuition waiver as well as receiving a certain amount of stipend provided by FSPs. Thus, we include the dummy variable for FSP recipients, who are all girls, in the conditional expenditure and core share equations. The regression results are reported in first three columns of Table 11. The coefficients on the girl dummy for the conditional expenditure and core share equations become more negative. The point estimates on the FSP dummy are positive in the conditional expenditure equation while they are significantly negative in the core share equation for both years.

To understand where this impact occurs, we report in Table 20 in Appendix E the marginal effects using item-by-item Tobit regressions, which are similar to those reported in Table 10, but they are calculated for both girl and FSP dummies this time. This analysis shows that the FSPs recipients spend less on tuition as expected because the tuition is waived for FSP recipients. For home tutoring and materials, the FSP recipients are getting more than nonrecepients, though this does not offset the amount of shortfall in the tuition expenditure. Thus, the FSP recipients still get less in the core component. For the items in the peripheral component, the FSP recipients get a higher expenditure in most items, especially in uniform, meal, and transportation. Because there may be systematic difference between FSPs recipients and nonrecepients, we cannot make a causal inference, but our results suggest that FSP did not increase the core expenditure conditional on enrollment.

Next, we study the spillover effect of FSPs exploiting the variations in the intensity of FSPs across regions and ages. The FSPs intensity variable, GRR, is constructed by the ratio of recipients among girls at division-age level. We also include its interaction term with the girl dummy in all the regressions and report the estimation results in Columns (4) to (6) of Table 11. The results show that girls living in more FSP-intensive divisions (for their age) are more likely to be enrolled in school. This indicates

[15] HES 1995 does not contain the information on the FSP status. HIES 2010 data was also not used because FSPs was already terminated. It should also be noted that the HIES 2000 data appears to underrepresent the FSP recipients. Based on BANBEIS (2006), the ratio of the number of FSPs recipients to the number of female enrolled secondary school students is 86%, while the figure directly derived from the HIES 2000 data is 59%. Therefore, the interpretation of the results for the year 2000 requires some caution. This issue does not exist for the year 2005.

Table 11: Three-part model estimation with the FSP status

| Year | *Coef.* | *d* | Cond *y* | Cond *s* | *d* | Cond *y* | Cond *s* |
|---|---|---|---|---|---|---|---|
| | | (1) | (2) | (3) | (4) | (5) | (6) |
| | Girl | 0.339*** | -0.245*** | -0.062*** | 0.228** | -0.236*** | -0.018 |
| | | (0.039) | (0.054) | (0.019) | (0.091) | (0.085) | (0.028) |
| | FSP | | 0.123** | -0.034** | | 0.149*** | -0.037** |
| **2000** | | | (0.049) | (0.015) | | (0.051) | (0.017) |
| | GRR | | | | 0.769** | -1.299*** | 0.247** |
| | | | | | (0.346) | (0.297) | (0.121) |
| | Girl × GRR | | | | 0.378 | -0.100 | -0.138* |
| | | | | | (0.286) | (0.260) | (0.078) |
| | *Obs.* | | 4878 | | | 4878 | |
| | Girl | 0.289*** | -0.178*** | -0.058*** | 0.110 | -0.107 | -0.007 |
| | | (0.034) | (0.034) | (0.014) | (0.093) | (0.072) | (0.025) |
| | FSP | | 0.046 | -0.026*** | | 0.075** | -0.025*** |
| **2005** | | | (0.036) | (0.009) | | (0.036) | (0.010) |
| | GRR | | | | 0.470 | -1.004*** | 0.020 |
| | | | | | (0.306) | (0.227) | (0.093) |
| | Girl × GRR | | | | 0.656** | -0.308 | -0.184** |
| | | | | | (0.315) | (0.233) | (0.081) |
| | *Obs.* | | 5638 | | | 5638 | |

Note: ***,**,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at the household level are reported in parentheses. GRR stands for the ratio of girl recipients to all girls at the division-age level. Additional controls include the covariates discussed in Table 4.

that FSPs may have a positive spillover effect on families living in the same area such that parents are more likely to enroll their children, particularly daughters, in school. However, there is no evidence that FSPs facilitate parental investment in the quality of education for girls. The coefficient on the interaction terms in the conditional education expenditure is negative for both 2000 and 2005, and the same coefficient in the conditional core share equation is significantly negative in both years.

### Impact on School Enrollment

The analysis so far has been based on a sample of children in their secondary-school age group at the time of survey. It is also possible to estimate the impact of FSPs by looking at the education from a larger sample, including adults at the time of survey with retrospective panel data. However, because of the lack of data on past education expenditure, we necessarily need to restrict our attention only to the enrollment. It is nevertheless still useful to verify whether the FSPs positively affected the enrollment from a different perspective.

To be specific, we use an empirical strategy similar to Heath and Mobarak (2015), where the enrollment status of individuals are retrospectively constructed for each year between 1960 and 2005 (for HIES 2005) or 2007 (for HIES 2010) from the highest completed grade and birth year. Then, we restrict the sample to the set of people who are aged between 6 to 15 in each year to match our school age definition.[16] Using these data, the impact of FSPs is identified by a triple difference estimator, where the differences are taken i) between boys and girls, ii) between the periods before and after 1994, the year in which the FSPs were rolled out nationwide, and iii) between those who have and those who have not completed primary schooling in a given year. The primary completion status essentially serves as the potential FSP-eligibility, and this is where our approach diverges from Heath and Mobarak (2015) and explains why our results qualitatively differ from theirs. We elaborate this point in Appendix D and argue that our specification is better suited to capture the impact of FSPs.

To be more specific, we estimate for girl $i$ in household $h$ in year $t$:

$$\begin{aligned} Enroll_{iht} &= \alpha_1 Girl_{ih} + \alpha_2 Post1994_t + \alpha_3 Primary\ Grad_{iht} + \alpha_4 Post1994_t \times Primary\ Grad_{iht} \\ &+ \alpha_5 Girl_{ih} \times Post1994_t + \alpha_6 Primary\ Grad_{iht} \times Girl_{ih} \\ &+ \alpha_7 Post1994_t \times Primary\ Grad_{iht} \times Girl_{ih} + \lambda_t^0 + \lambda_t^1 \times Girl_{ih} + \sum_{a=6}^{a=15} \beta_a^0 \times \mathbf{1}(Age = a) \\ &+ \sum_{a=6}^{a=15} \beta_a^1 \times \mathbf{1}(Age = a) \times Girl_{ih} + \theta_h + \epsilon_{iht}, \end{aligned} \quad (5)$$

[16] We also conduct similar analysis for those aged between 5 and 18 to match the definition of school-age children used in Heath and Mobarak (2015). The results are qualitatively the same.

where $\beta$s, $\lambda$s, and $\theta_h$ represent, respectively, age-gender-, time-gender-, and household-specific fixed effects.

The estimated coefficients of $\alpha_1$-$\alpha_7$ are presented in Table 12, where $\alpha_7$ is of our main interest. The table shows that the triple-difference estimates of the impact of FSPs on enrollment are all positive and significant. This is true whether or not we include the household fixed effects and whether we use the HIES data for the year 2005 or 2010. Therefore, based on our preferred specification of fixed effects model, the FSPs have increased the enrollment by 8.3-11.7 percentage points, which is broadly consistent with the marginal effects found in Table 9.[17] Therefore, while Table 12 has nothing to say about the effect of the FSPs on the educational expenditure or core share, it strengthens our finding that the FSPs had a positive impact on the secondary-school enrollment for girls and contributed to the observed reversal of the gender gap in the secondary enrollment.

Table 12: Triple-difference estimation of the impact of the FSPs on school enrollment with retrospective panel data

| | 2005 HIES (1) | 2010 HIES (2) | 2005 HIES (3) | 2010 HIES (4) |
|---|---|---|---|---|
| Girl | -0.205*** | -0.215*** | -0.211*** | -0.184*** |
| | (0.018) | (0.013) | (0.021) | (0.014) |
| Post 1994 | -0.378*** | 0.323*** | -0.439*** | 0.202*** |
| | (0.013) | (0.011) | (0.014) | (0.012) |
| Primary Grad | 0.164*** | 0.225*** | 0.625*** | 0.714*** |
| | (0.008) | (0.006) | (0.006) | (0.006) |
| Post 1994 × Primary Grad | -0.003 | -0.106*** | -0.051*** | -0.205*** |
| | (0.008) | (0.008) | (0.009) | (0.009) |
| Post 1994 × Girl | 0.230*** | 0.279*** | 0.220*** | 0.245*** |
| | (0.021) | (0.016) | (0.023) | (0.016) |
| Primary Grad × Girl | -0.144*** | -0.176*** | -0.125*** | -0.157*** |
| | (0.009) | (0.008) | (0.008) | (0.008) |
| Post 1994 × Primary Grad × Girl | 0.117*** | 0.083*** | 0.074*** | 0.042*** |
| | (0.011) | (0.011) | (0.013) | (0.013) |
| HH fixed effects | Y | Y | N | N |
| *Obs* | 236,258 | 348,516 | 236,258 | 348,516 |
| $R^2$ | 0.334 | 0.339 | 0.335 | 0.339 |

Note: ***,**,* denote statistical significance at 1, 5, 10 percent levels. The estimation is based on the OLS estimation of eq. (5). Standard errors clustered at the household level are reported in parentheses. Primary grad dummy is defined to have value 1 if the person has completed at least grade 5. The age group used in the analysis is 6 to 15.

[17] This comparison is not meant to be strict. Recall that Table 9 reports the marginal effect at the sample mean and not average marginal effect. In addition, the dependent variable for column (1) in Table 9 is probability of enrolled in secondary school, while we study the probability of remaining in school (regardless of the grade) in Table 12.

### Muting the FSPs Tuition Waiver

As mentioned above, the tuition waiver is an important instrument of the FSPs. The tuition waiver encourages enrollment but also tends to reduce the conditional expenditure and core share among the school enrollees. However, the latter negative effects may be spurious. This may be perhaps simply because the FSPs are crowding out the household's tuition expenditure for girls through tuition waiver; FSPs might not have any impact on the conditional expenditure and core share without the tuition waiver.

Table 13: Three-part model estimation with the impact of the tuition waiver muted

| Year | Model | $d$ | Cond $y$ | Cond $s$ |
|---|---|---|---|---|
| | Exclusion | 0.322*** | -0.081* | -0.062*** |
| | | (0.039) | (0.045) | (0.013) |
| **2000** | Imputation | 0.324*** | -0.072 | -0.055*** |
| | | (0.039) | (0.047) | (0.011) |
| | Baseline | 0.339*** | -0.174*** | -0.082*** |
| | | (0.039) | (0.049) | (0.014) |
| | Exclusion | 0.274*** | -0.079*** | -0.058*** |
| | | (0.035) | (0.028) | (0.011) |
| **2005** | Imputation | 0.279*** | -0.106*** | -0.050*** |
| | | (0.035) | (0.028) | (0.010) |
| | Baseline | 0.291*** | -0.154*** | -0.071*** |
| | | (0.034) | (0.027) | (0.012) |

Note: ***,**,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at the household level are reported in parentheses. Additional controls include the covariates discussed in Table 4. In the exclusion exercise, tuition fee is excluded from both total education expenditure and core expenditure. In the imputation exercise, we impute the tuition fee for FSP recipients using the predicted value from a linear model with survey year, female, grade, school type, enrollment, stipend recipient dummy, and district of residence. The baseline results are those presented in Table 4.

To see if this is a possible explanation, we attempt to mute the impact of the tuition waiver by conducting two alternative empirical exercises: exclusion and imputation. In the exclusion exercise, we remove the tuition fee from the calculations of both the total education expenditure and core expenditure. In the imputation exercise, we impute the tuition fee for the FSP recipients using the prediction model based on the survey year, gender, grade, school type, enrollment, stipend recipient dummy, and district information. The imputed tuition fee would reflect the tuition fee parents would have to spend had their daughter not received a tuition waiver. The results of these two exercises are presented in Table 13 together with the baseline model estimations for the ease of comparison.

Compared with the baseline results, the absolute value of the coefficient on the girl dummy gets smaller for all three equations after turning off the impact of tuition waiver either by exclusion or imputation. This finding indicates that a part of our finding is indeed driven by the spurious effect coming from the tuition waiver. However, as Table 13 shows, the signs and statistical significance of the coefficient on the girl dummy remain the same. Therefore, the earlier finding of the inconsistency in the direction of bias still remains valid even after muting the effects of tuition waiver.

### Impact on Timely Graduation from Secondary School

The FSPs appear to have improved the quantity of education in terms of enrollment. However, the FSPs have also been criticized for the lack of attention on education quality (Mahmud, 2003; Raynor and Wesson, 2006). Our analysis presented so far indeed aligns with this argument. If the quality of education for girls lags behind that for boys, the educational outcome is likely to suffer. In fact, we argue that the educational outcome should be the ultimate concern for researchers and policymakers.

Unfortunately, our data do not contain standard outcome measures of education such as test scores. Therefore, we use the completion of secondary school (roughly) on time as an indicator of education outcome. This is a reasonable outcome indicator because it is not trivial to pass the SSC exam.[18] Based on our age group classification, a child is regarded to have completed secondary school (roughly) on time if he/she has already passed at least grade 10 (SSC or equivalent) when he/she is in age 16-20. Because the HES for the year 1991 also contains information to construct the indicator for completion on time (but not individual-level information on education expenditure), we use five rounds of survey data for this analysis.

The year-by-year estimated effects of being a girl on the completion on time by OLS regressions are reported in columns (1)-(5) of Panel A, Table 14. It has became less pro-male and the beginning of the narrowing of the gap roughly corresponds to the onset of the FSPs, which seems to indicate that FSPs helped in closing the gender gap in timely completion of secondary education.

However, if we restrict the sample to those who have already completed primary education, the picture looks different as Panel B of Table 14 first five columns show. The gender gap in the timely completion of secondary education conditional on the completion of primary education is larger than in the unconditional sample except for the year 1991, which is before the start of the FSPs. This result indicate that the narrowing gender gap observed in Panel A may be due to the improvement in

[18]To complete secondary education, the child has to pass the SSC exam. As shown in Figure 1, the passing rate varies and may be as low as 40 percent depending on the year. Thus, without a certain level of mastery of secondary-level curriculum, the child will not be able to complete the secondary school.

girls' secondary enrollment. If more girls are enrolled, they have higher probability of completion. In addition, it is also consistent with our previous finding that the quality of education for girls conditional on enrollment consistently lagged behind that for boys.

Next, we conduct analysis with the FSP intensity (GRR; the ratio of recipients to all girls at the division-age level) and its interaction term with the girl dummy. Because the timely graduation from secondary school analysis reflects the accumulative impact of FSPs in the past 5 years, we use lagged GRR in the regression. That is, we use the GRR for the year 2000 [2005] in the analysis of timely graduation in the year 2005 [2010]. The results are presented in the last two columns of Table 14. For all children aged between 16 and 20, girls living in more program-intensive areas are less likely to graduate on time because the point estimates for the interaction term are negative. When we look only at the subsample of those who have completed primary education, the pro-male gender gap is significant in more FSPs intensive areas. Thus, we find no evidence that the FSPs improved girls' education quality.

## 7 Diagrammatical Analysis

While the enrollment rate for girls at the secondary level has substantially increased relative to that for boys over the last two decades in Bangladesh, the findings in the previous sections demonstrate that the quality and performance for girls lagged behind those for boys conditional on enrollment. Therefore, even though the FSPs appeared to have helped eliminate the gender gap in enrollment, it did not reduce or remove the gap in the investment in quality.

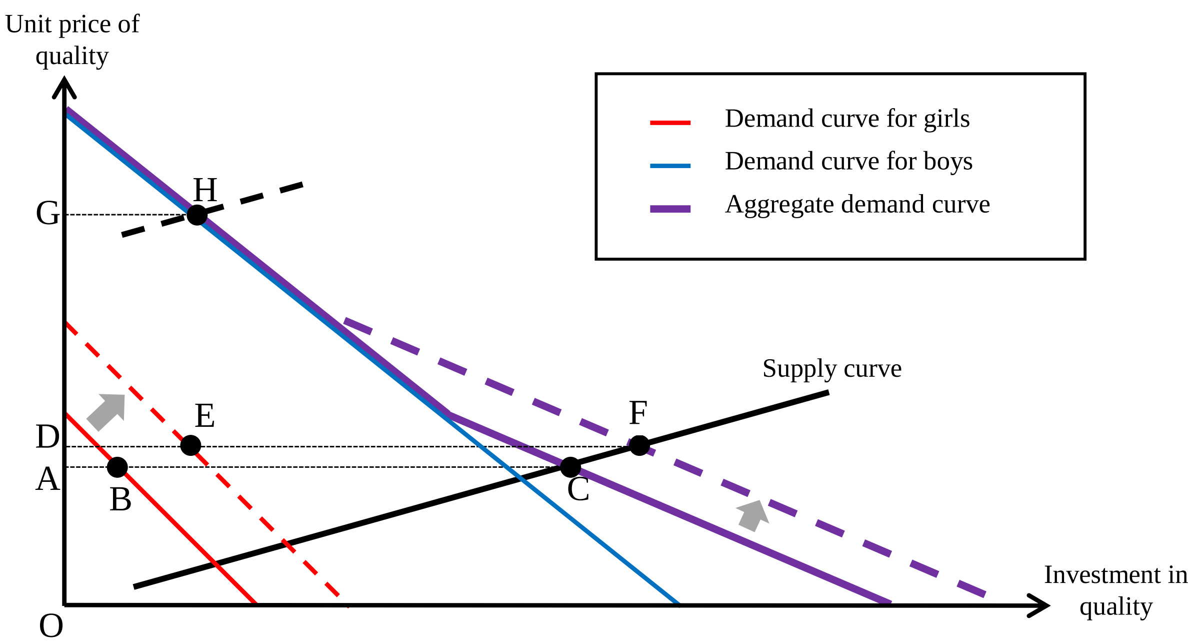


Figure 3: Demand and supply for the investment in the education quality.

Table 14: OLS regressions of completing Secondary education on time by year

| Sec complete on time<br>*Coef.* | **1991**<br>(1) | **1995**<br>(2) | **2000**<br>(3) | **2005**<br>(4) | **2010**<br>(5) | **2005**<br>(6) | **2010**<br>(7) |
|---|---|---|---|---|---|---|---|
| | | | *Panel A: All individuals aged 16-20* | | | | |
| Girl | -0.043*** | -0.053*** | -0.043*** | -0.014 | -0.005 | 0.004 | 0.065** |
| | (0.012) | (0.012) | (0.012) | (0.010) | (0.011) | (0.020) | (0.027) |
| Lagged GRR | | | | | | 0.242*** | 0.699*** |
| | | | | | | (0.091) | (0.091) |
| Girl×Lagged GRR | | | | | | -0.064 | -0.261*** |
| | | | | | | (0.070) | (0.094) |
| *Obs* | 3,043 | 3,752 | 3,988 | 5,055 | 5,316 | 5,055 | 5,316 |
| | | | *Panel B: All primary graduates aged 16-20* | | | | |
| Girl | -0.019 | -0.081*** | -0.063*** | -0.022* | -0.024* | 0.032 | 0.088*** |
| | (0.027) | (0.019) | (0.017) | (0.013) | (0.014) | (0.027) | (0.033) |
| Lagged GRR | | | | | | 0.345*** | 0.835*** |
| | | | | | | (0.122) | (0.115) |
| Girl×Lagged GRR | | | | | | -0.201** | -0.425*** |
| | | | | | | (0.094) | (0.116) |
| *Obs* | 1,223 | 2,113 | 2,621 | 3,716 | 4,089 | 3,716 | 4,089 |

Note: ***,** ,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at household level are reported in parentheses. The dependent variable is a dummy variable for the completion of secondary school on time, which takes one if an individual aged between 16 and 20 at the time of survey had already completed grade 10 or higher. Lagged GRR is the GRR at division-age level five years before the survey. In 2005 [2010], we use GRR for the year 2000 [2005]. In all regressions , the following covariates are also included : logarithmic expenditure per capita, logarithmic household size, the dummy variables for the household heads' education level (primary, secondary and higher), female head, wage-worker head, head's age and religion (muslim/hindu), and urban area. Panel A uses a sample of all individuals aged between 16 and 20 and Panel B uses a subsample of primary graduates among them.

A simple demand-supply diagram for the market of education quality for the secondary-school children is helpful for the understanding of this result. In Figure 3, we show the demand and supply curves for the market of education quality. While we abstract away from the details of the market of education quality, it can be considered as home tutoring for the ease of understanding. The black solid line in the figure represents the supply for the education quality. Its demand curves for girls and boys are shown in solid red and blue lines, respectively. In this figure, the demand for girls is always lower than that for boys, representing pro-male bias in the market for education quality. The aggregate demand is the kinked line in purple.

In this case, the equilibrium price is given by (the length of) OA and equilibrium demand for the boys and girls are AB and BC, respectively. The FSPs would reduce the cost of sending children to school, which may in turn increase the demand for the education quality, shifting the demand curve for the girls to the dashed red line. Then, the lower part of the aggregate demand curve will also shift to the dashed purple line. In this case, the new demand for girls will increase to DE whereas that for boys will decrease to EF. The demand for the education quality for boys decrease because of the higher equilibrium price resulting from higher completing demands from girls. In this picture, the aggregate investment in the education quality for girls would increase relative to that for boys as a result of the FSPs, and this is arguably the natural outcome that may be expected from the FSPs.

However, this is clearly inconsistent with our earlier empirical findings. Figure 3 also allows us to explain why the expected outcome may not occur. First, the demand for the quality of education may remain unchanged if the equilibrium price is above the choke price for girls. To demonstrate this point, suppose now that the supply curve is the black dashed line, such that the equilibrium price is GH regardless of the presence of FSP. In this case, the equilibrium demand for boys is GH and that for girls is zero, whether or not an FSP is in place. Second, it is also possible that the households who send girls to a secondary school because of the FSP may be unwilling to invest in the education quality. In this case, the demand curve for girls would be still the red solid line such that FSPs bring about no changes in the market for education quality.

These possibilities are also consistent with our finding that the pattern of gender bias is different in 1995 as shown in Table 4. Because the FSP coverage was substantially lower in 1995 than in 2000 and 2005, it is not surprising that the gender gap in enrollment was found insignificant in 1995. Furthermore, because many of the compliers—the girls who would go to school if they receive FPS but wouldn't go otherwise—are probably not covered in 1995, the effect on the core component is also small; if they were in school, they would receive very little investment in quality from their parents such that the estimated coefficient on the girl dummy for the core share would be small. While this discussion is somewhat speculative, it is consistent with both the empirical results and diagrammatic analysis presented above.

Incidentally, the diagram we use here is somewhat similar to Dang and Rogers (2015), who include private tutoring into the analysis of education in Vietnam. However, our model is distinct in two important aspects. First, the decision to enroll is not a particularly important decision in Vietnam as most children go to school and thus enrollment decision is not separately considered.[19] On the other hand, whether or not a child is sent to a secondary school remains an important household decision in Bangladesh. Second, we also explicitly distinguish between boys and girls and use the diagram to analyze the impact of a policy.

# 8 Conclusion

The allocation of educational resources is a key determinant of the human capital accumulation for children and their employment in later life. With affirmative action policies such as the FSPs in place, Bangladesh has achieved a significant improvement in the quantity measures of education such as school enrollment and years of schooling for girls.

Nevertheless, the quantity of education alone does not allow us to fully understand the gender gap in education. As we saw in Figure 1, girls underperformed boys in measures such as the passing rate and share of top students in the SSC exam. We hypothesized that this may be because the monetary investment in education, and the quality of education in particular, for girls may have been lower than that for boys. To test this hypothesis, we developed and estimated a three-part model to check the presence of gender-based gap in enrollment, total education expenditure, and expenditure share on the core component using four rounds of household survey data with detailed individual-level education expenditure data.

Our results demonstrate that there is a pro-female bias in enrollment but there is a pro-male bias in total education expenditure and core share decisions conditional on enrollment. This pattern is more pronounced at the secondary level than the primary level. We also find that the gender bias has been persistent and significant over the last one and half decades. While this pattern may be in part driven by the presence of the unobserved parental preference simultaneously affecting the family structure and intrahousehold allocation of education, this conclusion remains unchanged even if we alternatively estimate a linear model for each outcome with household fixed effects, which controls for all household-level observable and unobservable characteristics including the parental preference.

[19] Dang and Rogers (2015) report that 87 percent of children aged between 6 and 18 were enrolled in a school in the past 12 months in 2006.

We also performed item-by-item Tobit regressions. This analysis suggests that boys tend to receive more educational resources in the form of higher spending on tuition and home tutoring. While girls receive more resources for uniforms, this does not make up for the gap in other expenditure items such as tuition and home tutoring. Further, higher spending on uniforms is unlikely to lead to higher quality of education. Taken together, despite the parents' willingness to send girls to school, they appear to invest less in the quality of education for girls than for boys.

At a first glance, this pattern may appear puzzling because the direction of gender bias is inconsistent between the decision on enrollment and decision on educational expenditure and core share conditional on enrollment. This pattern is also puzzling because the neighboring countries such as India and Pakistan appear to have a clearer pro-male bias.

However, we have provided some evidence that this apparent inconsistency is at least in part driven by the FSPs. On one hand, the FSPs encourage more girls to attend secondary schools through stipend and tuition waivers. This finding that FSP had a positive impact on enrollment is robust. It holds across years and both in urban and rural areas. The conclusion remains unchanged when the triple-difference estimation is used with retrospective panel data instead of the three-part model. On the other hand, the FSPs do not incentivize parents to invest more in the quality of girls' education. While our finding appears to be partly driven by the presence of the tuition waiver in the FSPs, the significant pro-male bias in the educational expenditure and core share remains even after muting the effects of tuition waiver.

We further investigated the impact of the FSPs on timely graduation from secondary school as an outcome measure of education. This analysis indicates that the educational outcome for girls lags behind that for boys conditional on the completion of primary schooling. This is a result consistent with the aggregate measures presented in Figure 1 and also with the presence of pro-male bias in the education expenditure and core share. Therefore, putting these pieces of evidence together, we conclude that the FSPs were successful in bringing girls to classrooms but not in raising the quality of education they receive.

The current study highlights both the positive and (somewhat) negative aspects of FSPs. The positive news is that the FSPs did achieve what they intended to achieve. The FSPs certainly contributed to the promotion of gender equality in the secondary enrollment. However, the FSPs did not fully achieve gender equality in education because they did not raise the quality of education for girls relative to that for boys. There is no evidence that the FSPs attracted complementary investment from the households in the quality of education. Indeed, the exact opposite is true, if anything. Therefore, this study serves as a cautionary tale for policymakers in that the quantity of education does not necessarily tell the whole story about the gender gap in education. To truly achieve gender equality in Bangladesh, eliminating the gender gap in the investment of households would be needed.

Clearly, doing so is a challenge and requires us to know *why* there is a pro-male bias in the investment in quality of education, a question the current study is unable to answer. Still, it is possible to consider at least two rational reasons, besides the intrinsic gender bias of parents, why the pro-male bias may be observed. First, the returns from the investment in the quality of education may be different. If parents expect that girls will not be able to get a high-paying job even with high quality education, it would be rational for them not to invest in the quality of education. Note that they may still choose to send girls to school to receive the benefits of the FSPs. Second, similar to the first point, the parents may believe that they receive no returns in any case as girls will get married and bring about no economic returns to them. On the other hand, they may believe that boys are the important source of their retirement income. While exploring these possibilities is beyond the scope of this paper, we view this as an academically interesting question that has important policy implications to be explored in future research.

## References


Ahmed, M., A. K. Saleh, I. K. Nurul, and A. Romij (2007): *Access to education in Bangladesh: Country analytic review of primary and secondary education*, Brighton, UK and Dhaka, Bangladesh: Consortium for Research on Educational Access, Transitions and Equity (CREATE) and BRAC University Institute of Educational Development.

Aristei, D. and L. Pieroni (2008): "A double-hurdle approach to modelling tobacco consumption in Italy," *Applied Economics*, 40, 2463–2476.

Asadullah, M. N. (2006): "Returns to education in Bangladesh," *Education Economics*, 14, 453–468.

Asadullah, M. N. and N. Chaudhury (2009): "Reverse gender gap in schooling in Bangladesh: Insights from urban and rural households," *Journal of Development Studies*, 45, 1360–1380.

Aslam, M. and G. G. Kingdon (2008): "Gender and household education expenditure in Pakistan," *Applied Economics*, 40, 2573–2591.

Azam, M. and G. G. Kingdon (2013): "Are girls the fairer sex in India? Revisiting intra-household allocation of education expenditure," *World Development*, 42, 143–164.

BANBEIS (1995): "Bangladesh Educational Statistics," Dhaka: Bangladesh Bureau of Educational Information and Statistics (BANBEIS), Ministry of Education, the Government of Bangladesh.

——— (2006): "Bangladesh Educational Statistics," Dhaka: Bangladesh Bureau of Educational Information and Statistics (BANBEIS), Ministry of Education, the Government of Bangladesh.

——— (2010): "Bangladesh Educational Statistics," Dhaka: Bangladesh Bureau of Educational Information and Statistics (BANBEIS), Ministry of Education, the Government of Bangladesh.

Banerjee, A. and E. Duflo (2006): "Addressing absence," *Journal of Economic Perspectives*, 20, 117–132.

Bangladesh Bureau of Statistics (2002): "Statistical Yearbook of Bangladesh 2001-02," *Dhaka, Bangladesh.*

Bangladesh Ministry of Education (1996): "Stipend Operation Manual," Dhaka: Directorate of Secondary and Higher Education, Ministry of Education, Government of Bangladesh.

Begum, L., P. J. Grossman, and A. Islam (2016): "Gender Bias in Parental Attitude: An Experimental Approach," *Demography, forthcoming.*

Begum, L., A. Islam, and R. Smyth (2017): "Girl Power: Stipend Programs and the Education of Younger Siblings," *Journal of Development Studies*, 53, 1882–1898.

Behrman, J. A. (2015): "Do targeted stipend programs reduce gender and socioeconomic inequalities in schooling attainment? Insights from rural Bangladesh," *Demography*, 52, 1917–1927.

Behrman, J. R., R. A. Pollak, and P. Taubman (1982): "Parental preferences and provision for progeny," *Journal of Political Economy*, 90, 52–73.

Blunch, N.-H. and M. B. Das (2015): "Changing norms about gender inequality in education: Evidence from Bangladesh," *Demographic Research*, 32, 183–218.

Cragg, J. G. (1971): "Some statistical models for limited dependent variables with application to the demand for durable goods," *Econometrica*, 829–844.

Dang, H.-A. D. and F. Rogers (2015): "The Decision to Invest in Child Quality over QUantity: Household Size and Household Investment in Education in Vietnam," *World Bank Economic Review*, 30, 104–142.

Deaton, A. (1989): "Looking for boy-girl discrimination in household expenditure data," *World Bank Economic Review*, 3, 1–15.

Heath, R. and A. M. Mobarak (2015): "Manufacturing growth and the lives of Bangladeshi women," *Journal of Development Economics*, 115, 1–15.

Himaz, R. (2010): "Intrahousehold allocation of education expenditure: the case of Sri Lanka," *Economic Development and Cultural Change*, 58, 231–258.

Jensen, R. (2002): "Equal treatment, unequal outcomes," *Generating Sex Inequality Through Fertility Behaviour. John F. Kennedy School of Government, Harvard University.(Mimeographed.)*.

Jones, A. M. and S. T. Yen (2000): "A Box–Cox Double-hurdle Model," *Manchester School*, 68, 203–221.

Kenayathulla, H. B. (2016): "Gender differences in intra-household educational expenditures in Malaysia," *International Journal of Educational Development*, 46, 59–73.

Khandker, S., M. Pitt, and N. Fuwa (2003): "Subsidy to Promote Girls' Secondary Education: The Female Stipend Program in Bangladesh," MPRA Paper 23688, Munich Personal RePEc Archive.

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.

Lehmann, J.-Y., A. Nuevo-Chiquero, and M. Vidal-Fernández (2012): "Explaining the birth order effect: The role of prenatal and early childhood investments," IZA Discussion Paper 6755, IZA.

Li, D. and M. C. Tsang (2003): "Household decisions and gender inequality in education in rural China," *China: An International Journal*, 1, 224–248.

Mahmud, S. (2003): "Female secondary school stipend programme in Bangladesh: A critical assessment," Dhaka: Bangladesh Institute of Development Studies.

Majumder, A., A. Majumder, C. Mitra, and C. Mitra (2016): "Gender bias in household education expenditure: the case of West Bengal," *Indian Growth and Development Review*, 9, 129–150.

Masterson, T. (2012): "An empirical analysis of gender bias in education spending in Paraguay," *World Development*, 40, 583–593.

Raynor, J. and K. Wesson (2006): "The girls' stipend program in Bangladesh," *Journal of Education for International Development*, 2, 1–12.

Saha, A. (2013): "An assessment of gender discrimination in household expenditure on education in India," *Oxford Development Studies*, 41, 220–238.

Schurmann, A. T. (2009): "Review of the Bangladesh female secondary school stipend project using a social exclusion framework," *Journal of Health, Population and Nutrition*, 505–517.

Shonchoy, A. S. and M. Rabbani (2015): "The Bangladesh gender gap in education: Biased intra-household educational expenditures," IDE Discussion Paper 522, Institute of Developing Economies, Japan External Trade Organization (JETRO).

Wongmonta, S. and P. Glewwe (2017): "An analysis of gender differences in household education expenditure: the case of Thailand," *Education Economics*, 25, 183–204.

Yen, S. T. (1993): "Working wives and food away from home: the Box-Cox double hurdle model," *American Journal of Agricultural Economics*, 75, 884–895.

## References

Ahmed, M., A. K. Saleh, I. K. Nurul, and A. Romij (2007): *Access to education in Bangladesh: Country analytic review of primary and secondary education*, Brighton, UK and Dhaka, Bangladesh: Consortium for Research on Educational Access, Transitions and Equity (CREATE) and BRAC University Institute of Educational Development.

Aristei, D. and L. Pieroni (2008): "A double-hurdle approach to modelling tobacco consumption in Italy," *Applied Economics*, 40, 2463–2476.

Asadullah, M. N. (2006): "Returns to education in Bangladesh," *Education Economics*, 14, 453–468.

Asadullah, M. N. and N. Chaudhury (2009): "Reverse gender gap in schooling in Bangladesh: Insights from urban and rural households," *Journal of Development Studies*, 45, 1360–1380.

Aslam, M. and G. G. Kingdon (2008): "Gender and household education expenditure in Pakistan," *Applied Economics*, 40, 2573–2591.

Azam, M. and G. G. Kingdon (2013): "Are girls the fairer sex in India? Revisiting intra-household allocation of education expenditure," *World Development*, 42, 143–164.

BANBEIS (1995): "Bangladesh Educational Statistics," Dhaka: Bangladesh Bureau of Educational Information and Statistics (BANBEIS), Ministry of Education, the Government of Bangladesh.

——— (2006): "Bangladesh Educational Statistics," Dhaka: Bangladesh Bureau of Educational Information and Statistics (BANBEIS), Ministry of Education, the Government of Bangladesh.

——— (2010): "Bangladesh Educational Statistics," Dhaka: Bangladesh Bureau of Educational Information and Statistics (BANBEIS), Ministry of Education, the Government of Bangladesh.

Banerjee, A. and E. Duflo (2006): "Addressing absence," *Journal of Economic Perspectives*, 20, 117–132.

Bangladesh Bureau of Statistics (2002): "Statistical Yearbook of Bangladesh 2001-02," *Dhaka, Bangladesh.*

Bangladesh Ministry of Education (1996): "Stipend Operation Manual," Dhaka: Directorate of Secondary and Higher Education, Ministry of Education, Government of Bangladesh.

Begum, L., P. J. Grossman, and A. Islam (2016): "Gender Bias in Parental Attitude: An Experimental Approach," *Demography, forthcoming.*

Begum, L., A. Islam, and R. Smyth (2017): "Girl Power: Stipend Programs and the Education of Younger Siblings," *Journal of Development Studies*, 53, 1882–1898.

Behrman, J. A. (2015): "Do targeted stipend programs reduce gender and socioeconomic inequalities in schooling attainment? Insights from rural Bangladesh," *Demography*, 52, 1917–1927.

Behrman, J. R., R. A. Pollak, and P. Taubman (1982): "Parental preferences and provision for progeny," *Journal of Political Economy*, 90, 52–73.

Blunch, N.-H. and M. B. Das (2015): "Changing norms about gender inequality in education: Evidence from Bangladesh," *Demographic Research*, 32, 183–218.

Cragg, J. G. (1971): "Some statistical models for limited dependent variables with application to the demand for durable goods," *Econometrica*, 829–844.

Dang, H.-A. D. and F. Rogers (2015): "The Decision to Invest in Child Quality over QUantity: Household Size and Household Investment in Education in Vietnam," *World Bank Economic Review*, 30, 104–142.

Deaton, A. (1989): "Looking for boy-girl discrimination in household expenditure data," *World Bank Economic Review*, 3, 1–15.

Heath, R. and A. M. Mobarak (2015): "Manufacturing growth and the lives of Bangladeshi women," *Journal of Development Economics*, 115, 1–15.

Himaz, R. (2010): "Intrahousehold allocation of education expenditure: the case of Sri Lanka," *Economic Development and Cultural Change*, 58, 231–258.

Jensen, R. (2002): "Equal treatment, unequal outcomes," *Generating Sex Inequality Through Fertility Behaviour. John F. Kennedy School of Government, Harvard University.(Mimeographed.)*.

Jones, A. M. and S. T. Yen (2000): "A Box–Cox Double-hurdle Model," *Manchester School*, 68, 203–221.

Kenayathulla, H. B. (2016): "Gender differences in intra-household educational expenditures in Malaysia," *International Journal of Educational Development*, 46, 59–73.

Khandker, S., M. Pitt, and N. Fuwa (2003): "Subsidy to Promote Girls' Secondary Education: The Female Stipend Program in Bangladesh," MPRA Paper 23688, Munich Personal RePEc Archive.

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.

Lehmann, J.-Y., A. Nuevo-Chiquero, and M. Vidal-Fernández (2012): "Explaining the birth order effect: The role of prenatal and early childhood investments," IZA Discussion Paper 6755, IZA.

Li, D. and M. C. Tsang (2003): "Household decisions and gender inequality in education in rural China," *China: An International Journal*, 1, 224–248.

Mahmud, S. (2003): "Female secondary school stipend programme in Bangladesh: A critical assessment," Dhaka: Bangladesh Institute of Development Studies.

Majumder, A., A. Majumder, C. Mitra, and C. Mitra (2016): "Gender bias in household education expenditure: the case of West Bengal," *Indian Growth and Development Review*, 9, 129–150.

Masterson, T. (2012): "An empirical analysis of gender bias in education spending in Paraguay," *World Development*, 40, 583–593.

Raynor, J. and K. Wesson (2006): "The girls' stipend program in Bangladesh," *Journal of Education for International Development*, 2, 1–12.

Saha, A. (2013): "An assessment of gender discrimination in household expenditure on education in India," *Oxford Development Studies*, 41, 220–238.

Schurmann, A. T. (2009): "Review of the Bangladesh female secondary school stipend project using a social exclusion framework," *Journal of Health, Population and Nutrition*, 505–517.

Shonchoy, A. S. and M. Rabbani (2015): "The Bangladesh gender gap in education: Biased intra-household educational expenditures," IDE Discussion Paper 522, Institute of Developing Economies, Japan External Trade Organization (JETRO).

Wongmonta, S. and P. Glewwe (2017): "An analysis of gender differences in household education expenditure: the case of Thailand," *Education Economics*, 25, 183–204.

Yen, S. T. (1993): "Working wives and food away from home: the Box-Cox double hurdle model," *American Journal of Agricultural Economics*, 75, 884–895.

# A Derivation of the likelihood function for the three-part model

In total, there are four separate cases to consider to construct the likelihood function for the three-part model:

**Case 1:** $d = 0$.

$$l_1 = P(\epsilon_d \leq -x_d'\beta_d) = \Phi(-x_d'\beta_d).$$

**Case 2:** $d = 1, y = y, s = 0$.

$$\begin{aligned}
l_2 &= P(d = 1, y = y, s = 0) \\
&= \frac{1}{y}P(\epsilon_d > -x_d'\beta_d,\ \epsilon_y = \log(y) - x_y'\beta_y,\ \epsilon_s \leq -x_s'\beta_s) \\
&= \frac{1}{y}P(-\epsilon_d \leq x_d'\beta_d,\ \epsilon_s \leq -x_s'\beta_s|\ \epsilon_y = \log(y) - x_y'\beta_y)f(\log(y) - x_y'\beta_y),
\end{aligned}$$

where $f(\cdot)$ is the density function of $\epsilon_y$.

We rearrange the distribution of the error terms as follows:

$$\begin{bmatrix} -\epsilon_d \\ \epsilon_s \\ \epsilon_y \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} 1 & -\rho_{ds}\sigma_s & -\rho_{dy}\sigma_y \\ -\rho_{ds}\sigma_s & \sigma_s^2 & \rho_{ys}\sigma_y\sigma_s \\ -\rho_{dy}\sigma_y & \rho_{ys}\sigma_y\sigma_s & \sigma_y^2. \end{bmatrix}\right).$$

$(-\epsilon_d, \epsilon_s)^T$ given $\epsilon_y$ follows bivariate normal distribution with:

$$\mathbf{E}\left(\left(\begin{array}{c} -\epsilon_d \\ \epsilon_s \end{array}\right) \middle| \epsilon_y\right) = \left(\begin{array}{c} 0 \\ 0 \end{array}\right) + \left(\begin{array}{c} -\rho_{dy}\sigma_y \\ \rho_{ys}\sigma_y\sigma_s \end{array}\right)\frac{1}{\sigma_y^2}(\epsilon_y - 0) = \left(\begin{array}{c} -\frac{\rho_{dy}}{\sigma_y}\epsilon_y \\ \frac{\rho_{ys}\sigma_s}{\sigma_y}\epsilon_y \end{array}\right),$$

and

$$\begin{aligned}
\mathbf{Var}\left(\left(\begin{array}{c} -\epsilon_d \\ \epsilon_s \end{array}\right) \middle| \epsilon_y\right) &= \left(\begin{array}{cc} 1 & -\rho_{ds}\sigma_s \\ -\rho_{ds}\sigma_s & \sigma_s^2 \end{array}\right) - \left(\begin{array}{c} -\rho_{dy}\sigma_y \\ \rho_{ys}\sigma_y\sigma_s \end{array}\right)\frac{1}{\sigma_y^2}\left(\begin{array}{cc} -\rho_{dy}\sigma_y & \rho_{ys}\sigma_y\sigma_s \end{array}\right) \\
&= \left(\begin{array}{cc} 1 - \rho_{dy}^2 & (\rho_{dy}\rho_{ys} - \rho_{ds})\sigma_s \\ (\rho_{dy}\rho_{ys} - \rho_{ds})\sigma_s & (1 - \rho_{ys}^2)\sigma_s^2 \end{array}\right).
\end{aligned}$$

Then, we have:

$$\begin{aligned} & P(-\epsilon_d \leq x_d'\beta_d,\ \epsilon_s \leq -x_s'\beta_s |\ \epsilon_y = \log(y) - x_y'\beta_y) \\ = \quad & \Psi\left(\frac{x_d'\beta_d + \rho_{dy}\epsilon_y/\sigma_y}{\sqrt{1-\rho_{dy}^2}},\ -\frac{x_s'\beta_s + \rho_{ys}\sigma_s\epsilon_y/\sigma_y}{\sigma_s\sqrt{1-\rho_{ys}^2}},\ \frac{\rho_{dy}\rho_{ys} - \rho_{ds}}{\sqrt{(1-\rho_{dy}^2)(1-\rho_{ys}^2)}}\right), \end{aligned}$$

and

$$f(\log(y) - x_y'\beta_y) = \frac{1}{\sigma_y}\phi(\frac{\log(y) - x_y'\beta_y}{\sigma_y}).$$

Thus, the likelihood for this case is:

$$l_2 = \frac{\phi(e_y)}{y\sigma_y} \cdot \Psi\left(\frac{x_d'\beta_d + \rho_{dy}e_y}{\sqrt{1-\rho_{dy}^2}},\ -\frac{x_s'\beta_s + \rho_{ys}\sigma_s e_y}{\sigma_s\sqrt{1-\rho_{ys}^2}},\ \frac{\rho_{dy}\rho_{ys} - \rho_{ds}}{\sqrt{(1-\rho_{dy}^2)(1-\rho_{ys}^2)}}\right).$$

**Case 3:** $d = 1, y = y, s \in (0, 1)$.

$$\begin{aligned} l_3 &= P(d = 1, y = y, s = s) \\ &= \frac{1}{y}P(\epsilon_d > -x_d'\beta_d,\ \epsilon_y = \log(y) - x_y'\beta_y,\ \epsilon_s = s - x_s'\beta_s) \\ &= \frac{1}{y}P(-\epsilon_d \leq x_d'\beta_d |\ \epsilon_y = \log(y) - x_y'\beta_y,\ \epsilon_s = s - x_s'\beta_s)g(\log(y) - x_y'\beta_y,\ s - x_s'\beta_s), \end{aligned}$$

where $g(\cdot, \cdot)$ is the joint density function for $\epsilon_y$ and $\epsilon_s$.

Let the submatrix $\Sigma_{11}$ be

$$\Sigma_{11} = \begin{pmatrix} \sigma_y^2 & \rho_{ys}\sigma_y\sigma_s \\ \rho_{ys}\sigma_y\sigma_s & \sigma_s^2 \end{pmatrix}.$$

Thus, we have

$$\Sigma_{11}^{-1} = \frac{1}{(1-\rho_{ys}^2)\sigma_y^2\sigma_s^2} \begin{pmatrix} \sigma_s^2 & -\rho_{ys}\sigma_y\sigma_s \\ -\rho_{ys}\sigma_y\sigma_s & \sigma_y^2 \end{pmatrix},$$

where the determinant of $\Sigma_{11}$ is $|\Sigma_{11}| = (1 - \rho_{ys}^2)\sigma_y^2\sigma_s^2$.

$-\epsilon_d$ given $\epsilon_y$ and $\epsilon_s$ follows normal distribution with:

$$\begin{aligned}
\mathbf{E}\left(-\epsilon_d|\epsilon_y,\ \epsilon_s\right) &= 0+\frac{1}{|\Sigma_{11}|}\begin{pmatrix} -\rho_{dy}\sigma_y & -\rho_{ds}\sigma_s \end{pmatrix}\begin{pmatrix} \sigma_s^2 & -\rho_{ys}\sigma_y\sigma_s \\ -\rho_{ys}\sigma_y\sigma_s & \sigma_y^2 \end{pmatrix}\begin{pmatrix} \epsilon_y \\ \epsilon_s \end{pmatrix} \\
&= -\frac{1}{(1-\rho_{ys}^2)\sigma_y^2\sigma_s^2}\begin{pmatrix} (\rho_{dy}-\rho_{ds}\rho_{ys})\sigma_y\sigma_s^2 & (\rho_{ds}-\rho_{dy}\rho_{ys})\sigma_y^2\sigma_s) \end{pmatrix}\begin{pmatrix} \epsilon_y \\ \epsilon_s \end{pmatrix} \\
&= -\frac{(\rho_{dy}-\rho_{ds}\rho_{ys})\sigma_s\epsilon_y+(\rho_{ds}-\rho_{dy}\rho_{ys})\sigma_y\epsilon_s}{(1-\rho_{ys}^2)\sigma_y\sigma_s},
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{Var}\left(-\epsilon_d|\epsilon_y,\ \epsilon_s\right) &= 1-\frac{1}{|\Sigma_{11}|}\begin{pmatrix} -\rho_{dy}\sigma_y & -\rho_{ds}\sigma_s \end{pmatrix}\begin{pmatrix} \sigma_s^2 & -\rho_{ys}\sigma_y\sigma_s \\ -\rho_{ys}\sigma_y\sigma_s & \sigma_y^2 \end{pmatrix}\begin{pmatrix} -\rho_{dy}\sigma_y \\ -\rho_{ds}\sigma_s \end{pmatrix} \\
&= 1-\frac{1}{(1-\rho_{ys}^2)\sigma_y^2\sigma_s^2}\begin{pmatrix} -(\rho_{dy}-\rho_{ds}\rho_{ys})\sigma_y\sigma_s^2 & -(\rho_{ds}-\rho_{dy}\rho_{ys})\sigma_y^2\sigma_s \end{pmatrix}\begin{pmatrix} -\rho_{dy}\sigma_y \\ -\rho_{ds}\sigma_s \end{pmatrix} \\
&= 1-\frac{(\rho_{dy}-\rho_{ds}\rho_{ys})\rho_{dy}+(\rho_{ds}-\rho_{dy}\rho_{ys})\rho_{ds}}{(1-\rho_{ys}^2)} \\
&= \frac{1-\rho_{ys}^2-\rho_{dy}^2-\rho_{ds}^2+2\rho_{dy}\rho_{ds}\rho_{ys}}{1-\rho_{ys}^2}.
\end{aligned}$$

We then have

$$\begin{aligned}
& P(-\epsilon_d \le x_d'\beta_d|\ \epsilon_y=\log(y)-x_y'\beta_y,\ \epsilon_s=s-x_s'\beta_s) \\
= \ & \Phi\left(\frac{x_d'\beta_d(1-\rho_{ys}^2)+(\rho_{dy}-\rho_{ds}\rho_{ys})(\log(y)-x_y'\beta_y)/\sigma_y+(\rho_{ds}-\rho_{dy}\rho_{ys})(s-x_s'\beta_s)/\sigma_s}{\sqrt{(1-\rho_{ys}^2-\rho_{dy}^2-\rho_{ds}^2+2\rho_{dy}\rho_{ds}\rho_{ys})(1-\rho_{ys}^2)}}\right),
\end{aligned}$$

and

$$
\begin{aligned}
& g(\epsilon_y, \epsilon_s) = g(\log(y) - x_y'\beta_y,\ s - x_s'\beta_s) \\
= \quad & \frac{1}{2\pi\sigma_y\sigma_s\sqrt{1-\rho_{ys}^2}} \exp\left[ -\frac{1}{2} \begin{pmatrix} \epsilon_y & \epsilon_s \end{pmatrix} \frac{1}{|\Sigma_{11}|} \begin{pmatrix} \sigma_s^2 & -\rho_{ys}\sigma_y\sigma_s \\ -\rho_{ys}\sigma_y\sigma_s & \sigma_y^2 \end{pmatrix} \begin{pmatrix} \epsilon_y \\ \epsilon_s \end{pmatrix} \right] \\
= \quad & \frac{1}{2\pi\sigma_y\sigma_s\sqrt{1-\rho_{ys}^2}} \exp\left[ -\frac{\epsilon_y^2\sigma_s^2 - 2\rho_{ys}\sigma_y\sigma_s\epsilon_y\epsilon_s + \epsilon_s^2\sigma_y^2}{2(1-\rho_{ys}^2)\sigma_y^2\sigma_s^2} \right] \\
= \quad & \frac{1}{\sigma_y\sigma_s\sqrt{1-\rho_{ys}^2}} \phi\left( \frac{\epsilon_y}{\sigma_y\sqrt{1-\rho_{ys}^2}} \right) \phi\left( \frac{\epsilon_s}{\sigma_s\sqrt{1-\rho_{ys}^2}} \right) \exp\left( \rho_{ys} \frac{\epsilon_y\epsilon_s}{(1-\rho_{ys}^2)\sigma_y\sigma_s} \right) \\
= \quad & \frac{1}{\sigma_y\sigma_s\sqrt{1-\rho_{ys}^2}} \phi\left( \frac{\log(y) - x_y'\beta_y}{\sigma_y\sqrt{1-\rho_{ys}^2}} \right) \phi\left( \frac{s - x_s'\beta_s}{\sigma_s\sqrt{1-\rho_{ys}^2}} \right) \exp\left( \rho_{ys} \frac{(\log(y) - x_y'\beta_y)(s - x_s'\beta_s)}{(1-\rho_{ys}^2)\sigma_y\sigma_s} \right).
\end{aligned}
$$

Thus, the likelihood for this case is:

$$
\begin{aligned}
l_3 \quad = \quad & \frac{1}{y\sigma_y\sigma_s\sqrt{1-\rho_{ys}^2}} \Phi\left( \frac{x_d'\beta_d(1-\rho_{ys}^2) + (\rho_{dy} - \rho_{ds}\rho_{ys})e_y + (\rho_{ds} - \rho_{dy}\rho_{ys})e_s}{\sqrt{(1-\rho_{ys}^2 - \rho_{dy}^2 - \rho_{ds}^2 + 2\rho_{dy}\rho_{ds}\rho_{ys})(1-\rho_{ys}^2)}} \right) \\
& \cdot\phi\left( \frac{e_y}{\sqrt{1-\rho_{ys}^2}} \right) \phi\left( \frac{e_s}{\sqrt{1-\rho_{ys}^2}} \right) \exp\left( \rho_{ys} \frac{e_y e_s}{1-\rho_{ys}^2} \right).
\end{aligned}
$$

**Case 4:** $d = 1, y = y, s = 1$.

$$
\begin{aligned}
l_4 \quad &= \quad P(d = 1, y = y, s = 1) \\
&= \quad \frac{1}{y} P(\epsilon_d > -x_d'\beta_d,\ \epsilon_y = \log(y) - x_y'\beta_y,\ \epsilon_s \geq 1 - x_s'\beta_s) \\
&= \quad \frac{1}{y} P(-\epsilon_d \leq x_d'\beta_d,\ -\epsilon_s \leq x_s'\beta_s - 1|\ \epsilon_y = \log(y) - x_y'\beta_y) f(\log(y) - x_y'\beta_y)
\end{aligned}
$$

We rearrange the distribution of the error terms as follows:

$$
\begin{bmatrix} -\epsilon_d \\ -\epsilon_s \\ \epsilon_y \end{bmatrix} \sim N\left( \mathbf{0}, \begin{bmatrix} 1 & \rho_{ds}\sigma_s & -\rho_{dy}\sigma_y \\ \rho_{ds}\sigma_s & \sigma_s^2 & -\rho_{ys}\sigma_y\sigma_s \\ -\rho_{dy}\sigma_y & -\rho_{ys}\sigma_y\sigma_s & \sigma_y^2 \end{bmatrix} \right).
$$

$(-\epsilon_d, -\epsilon_s)^T$ given $\epsilon_y$ follows bivariate normal distribution with:

$$
\mathbf{E}\left( \left. \begin{pmatrix} -\epsilon_d \\ -\epsilon_s \end{pmatrix} \right| \epsilon_y \right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -\rho_{dy}\sigma_y \\ -\rho_{ys}\sigma_y\sigma_s \end{pmatrix} \frac{1}{\sigma_y^2}(\epsilon_y - 0) = \begin{pmatrix} -\frac{\rho_{dy}}{\sigma_y}\epsilon_y \\ -\frac{\rho_{ys}\sigma_s}{\sigma_y}\epsilon_y \end{pmatrix},
$$

and

$$\mathbf{Var}\left(\left.\begin{pmatrix} -\epsilon_d \\ -\epsilon_s \end{pmatrix}\right| \epsilon_y\right) = \begin{pmatrix} 1 & \rho_{ds}\sigma_s \\ \rho_{ds}\sigma_s & \sigma_s^2 \end{pmatrix} - \begin{pmatrix} -\rho_{dy}\sigma_y \\ -\rho_{ys}\sigma_y\sigma_s \end{pmatrix} \frac{1}{\sigma_y^2} \begin{pmatrix} -\rho_{dy}\sigma_y & -\rho_{ys}\sigma_y\sigma_s \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \rho_{ds}\sigma_s \\ \rho_{ds}\sigma_s & \sigma_s^2 \end{pmatrix} - \begin{pmatrix} \rho_{dy}^2 & \rho_{dy}\rho_{ys}\sigma_s \\ \rho_{dy}\rho_{ys}\sigma_s & \rho_{ys}^2\sigma_s^2 \end{pmatrix}$$

$$= \begin{pmatrix} 1-\rho_{dy}^2 & (\rho_{ds}-\rho_{dy}\rho_{ys})\sigma_s \\ (\rho_{ds}-\rho_{dy}\rho_{ys})\sigma_s & (1-\rho_{ys}^2)\sigma_s^2 \end{pmatrix}.$$

Then, we have

$$P(-\epsilon_d \leq x_d'\beta_d,\ -\epsilon_s \leq x_s'\beta_s - 1|\ \epsilon_y = \log(y) - x_y'\beta_y)$$

$$= \Psi\left(\frac{x_d'\beta_d + \rho_{dy}(\log(y) - x_y'\beta_y)/\sigma_y}{\sqrt{1-\rho_{dy}^2}},\ \frac{x_s'\beta_s - 1 + \rho_{ys}\sigma_s(\log(y) - x_y'\beta_y)/\sigma_y}{\sigma_s\sqrt{1-\rho_{ys}^2}},\ \frac{\rho_{ds}-\rho_{dy}\rho_{ys}}{\sqrt{(1-\rho_{dy}^2)(1-\rho_{ys}^2)}}\right),$$

and

$$f(\log(y) - x_y'\beta_y) = \frac{1}{\sigma_y}\phi\left(\frac{\log(y) - x_y'\beta_y}{\sigma_y}\right).$$

Thus, the likelihood for this case is:

$$l_4 = \frac{\phi(e_y)}{y\sigma_y}\cdot\Psi\left(\frac{x_d'\beta_d + \rho_{dy}e_y}{\sqrt{1-\rho_{dy}^2}},\ \frac{x_s'\beta_s - 1 + \rho_{ys}\sigma_s e_y}{\sigma_s\sqrt{1-\rho_{ys}^2}},\ \frac{\rho_{ds}-\rho_{dy}\rho_{ys}}{\sqrt{(1-\rho_{dy}^2)(1-\rho_{ys}^2)}}\right),$$

where $e_y = \frac{\log(y) - x_y'\beta_y}{\sigma_y}$ and $e_s = \frac{s - x_s'\beta_s}{\sigma_s}$.

## B Derivation of marginal effects

The equation for the expected enrollment is straightforward. The equation for the conditional expenditure can be derived as follows:

$$\begin{aligned}
E(y|d=1) &= \int_0^\infty yf(y|d=1)dy = \int_0^\infty yf(y|\epsilon_1 > -x_d'\beta_d)dy \\
&= \int_0^\infty y\frac{1}{y}f(\epsilon_y|\epsilon_1 > -x_d'\beta_d)dy = \int_0^\infty \frac{f(\epsilon_y, \epsilon_1 > -x_d'\beta_d)}{P(\epsilon_1 > -x_d'\beta_d)}dy \\
&= \int_0^\infty \frac{f(\epsilon_1 > -x_d'\beta_d|\epsilon_y)f(\epsilon_y)}{P(\epsilon_1 > -x_d'\beta_d)}dy \\
&= \int_0^\infty \frac{\Phi\left(\frac{x_d'\beta_d+\rho_{dy}\epsilon_y/\sigma_y}{\sqrt{1-\rho_{dy}^2}}\right)\phi\left(\frac{\epsilon_y}{\sigma_y}\right)/\sigma_y}{\Phi(x_d'\beta_d)}dy,
\end{aligned}$$

where $\epsilon_y = \log(y) - x_y'\beta_y$.

The unconditional expectation of $y$ is:

$$E(y) = P(d=1)E(y|d=1) = \int_0^\infty \frac{1}{\sigma_y}\Phi\left(\frac{x_d'\beta_d + \rho_{dy}\epsilon_y/\sigma_y}{\sqrt{1-\rho_{dy}^2}}\right)\phi\left(\frac{\epsilon_y}{\sigma_y}\right)dy.$$

Unconditional expectation of the core expenditure $ys$ is:

$$\begin{aligned}
E(ys) &= \int_0^1\int_0^\infty ysf(y,s)dyds = \int_0^\infty yf(y|\epsilon_1 > -x_d'\beta_d)dy \\
&= \int_0^\infty y\cdot 1\cdot f(d=1,y,s=1)dy + \int_0^1\int_0^\infty ysf(d=1,y,s)dyds \\
&= \int_0^\infty \frac{1}{\sigma_y}\phi(\frac{\epsilon_y}{\sigma_y})\Psi\left(\frac{x_d'\beta_d + \rho_{dy}\epsilon_y/\sigma_y}{\sqrt{1-\rho_{dy}^2}}, \frac{x_s'\beta_s - 1 + \rho_{ys}\sigma_s\epsilon_y/\sigma_y}{\sigma_s\sqrt{1-\rho_{ys}^2}}, \frac{\rho_{ds}-\rho_{dy}\rho_{ys}}{\sqrt{(1-\rho_{dy}^2)(1-\rho_{ys}^2)}}\right)dy \\
&+ \int_0^1\int_0^\infty ys\frac{1}{y\sigma_y\sigma_s\sqrt{1-\rho_{ys}^2}}\Phi\left(\frac{x_d'\beta_d(1-\rho_{ys}^2) + (\rho_{dy}-\rho_{ds}\rho_{ys})\epsilon_y/\sigma_y + (\rho_{ds}-\rho_{dy}\rho_{ys})\epsilon_s/\sigma_s}{\sqrt{(1-\rho_{ys}^2-\rho_{dy}^2-\rho_{ds}^2+2\rho_{dy}\rho_{ds}\rho_{ys})(1-\rho_{ys}^2)}}\right) \\
&\times\phi\left(\frac{\epsilon_y}{\sigma_y\sqrt{1-\rho_{ys}^2}}\right)\phi\left(\frac{\epsilon_s}{\sigma_s\sqrt{1-\rho_{ys}^2}}\right)\exp\left(\rho_{ys}\frac{\epsilon_y\epsilon_s}{\sigma_y\sigma_s(1-\rho_{ys}^2)}\right)dyds,
\end{aligned}$$

where $\epsilon_s = s - x_s'\beta_s$.

The conditional (on positive educational expenditure) expectation of the core expenditure is:

$$E(ys|d=1) = E(ys|d=1) = \frac{E(ys)}{P(d=1)} = \frac{E(ys)}{\Phi\left(x_d'\beta_d\right)}.$$

For each of these equations, we can compute the sample analogues by replacing the parameters ($\theta$) with the ML estimates ($\hat{\theta}_{ML}$) given covariates.

We obtain the standard errors for the marginal effects by simulation. That is, we first draw the parameter $\theta$ from a multivariate normal distribution, where its mean and variance respectively follow the point estimate and its variance-covariance matrix from the ML estimation. We then calculate the marginal effects again with the drawn value of $\theta$ using the expressions above. By repeating this 100 times and taking the standard deviation of the estimates of the marginal effect, we obtain a standard error.

In principle, we can calculate the marginal effect for each observation and then calculate the average marginal effect over all observations. However, we choose to calculate only the marginal effects at the sample mean, where the sample mean of the whole sample [subsample of secondary-school enrollees] is used for the marginal effects on the probability of enrollment and unconditional quantities [conditional quantities] to mitigate the computational burden.[20]

# C Tuition and quality of education

To understand the relationship between the quality of education and tuition fee, we would ideally run a regression of tuition fee on an indicator of education quality. However, we do not have school-level data that can be linked to HES/HIES data. Instead, we run a regression of the average test score on the average tuition per student at the school level using the dataset for the Comparing Food versus Cash for Education (FFE-CFE) program for the years 2000 and 2003 collected by the International Food Policy Research Institute. While these data are available only for primary schools, this is the only data set that allows us to link the tuition fee and educational outcome in Bangladesh.

As Figure 4 shows, the average test score is higher in schools that impose a higher tuition both in 2000 and 2003. Clearly, this does not serve as definitive evidence that higher tuition reflects higher educational quality at the secondary level for a number of reasons. First, the data we use here are for the primary level and not the secondary level. Second, the data are not nationally representative and the sample selection may be an issue. Third, we do not consider the effect of endogenous school choice; it may be the case that those children with parents who can afford to pay a high tuition are those with high innate ability or those who receive complementary home tutoring. Nevertheless, the evidence that Figure 4 provides is consistent with the possibility that higher tuition reflects higher educational quality.

[20] Matlab was used for computation of the marginal effects and STATA was used in the rest of the analysis.

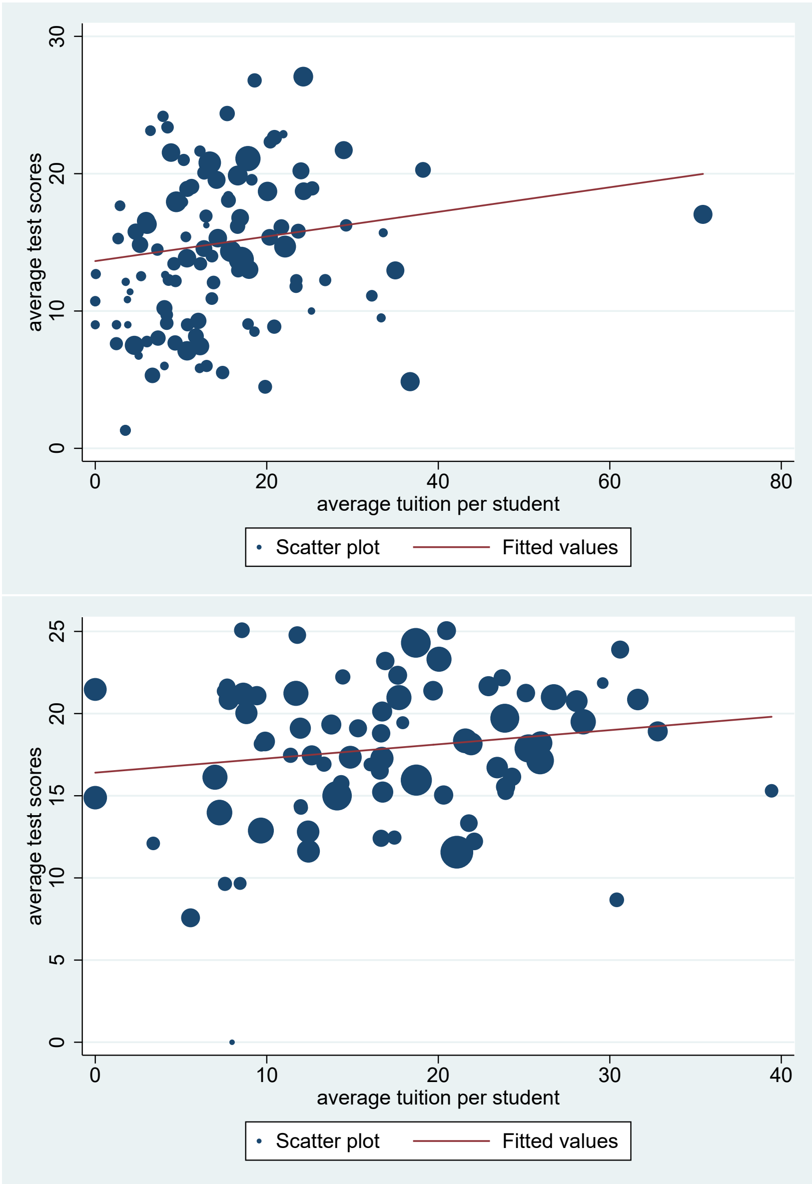


Figure 4: The scatter plot of the average test score and the average tuition fee charged by the school for the years 2000 (top) and 2003 (bottom) based on the FFE-CFE data. The size of each dot is proportionate to the number of enrolled students in the school and the red line represents the linear fit (weighted). The slope of the line are significantly different from 0 at a 10 percent level.

It should be noted that the tuition fee is not a simple reflection of school type. In the FFE-CFE data, there is indeed a substantial variation in the tuition per student both in private and public schools. Correspondingly, there appear to be significant variations in quality within each type. Casual observations of schools indicate that most of top schools are private in Bangladesh. On the other hand, the BANBEIS database suggests that private schools are not only smaller than public schools on average but also their quality is lower. For example, the average quality of teachers in private schools is worse than that for public schools as measured by the fraction of trained teachers. Student-teacher ratios for private schools were, if anything, slightly higher than those for public schools at the secondary level in the past, even though they are very similar today.[21] Therefore, the average quality of private schools appears to be lower than that of public schools.

## D Comparison with Heath and Mobarak (2015)

Heath and Mobarak (2015) find no evidence that the FSPs make girls more likely to remain in school compared with boys. In fact, their triple-difference estimate of the impact of FSPs is negative (Table 9 of Heath and Mobarak (2015)). However, our analysis differs in three important aspects.

First, the data in the current study are national representative and were collected in 2010, while their data were collected only from Dhaka and Gazipur districts in 2009. Second, the information available in the datasets are different. Heath and Mobarak (2015) use the school entry age, timing and length of any interruptions in schooling, and years of completed education to construct the grade level and enrollment status for each year. The HIES data we use only contains completed grade and age information. Therefore, we need to make additional assumptions that all children start schooling at age 6 and that there is no grade repetition. For example, consider a woman who had completed grade 7 and was 30 years of age in 2010. Given this information, we infer that she started schooling at age 6 in 1986 and remained enrolled in school until 1992. Based on this inference, we also know that she was a primary graduate (post5 as shown in the analysis below) from 1991 onwards.[22] After constructing the history of enrollment, we use only the observations that correspond to the individuals aged between 5 and 18 in each year to match the sample used in Heath and Mobarak (2015). The results are similar even when the cohort aged between 6 and 15 is used each year as in our main context.

Finally and most importantly, we take the (third) difference between those who graduated from primary school—completed grade 5 ("post5")—and those who did not after taking the differences

[21] This is not true at the primary level. Private schools are smaller and student-teacher ratios in private schools are much lower than those in public schools.

[22] Note that the girl was enrolled in grade 6 in 1991 and she had completed grade 5 in that year.

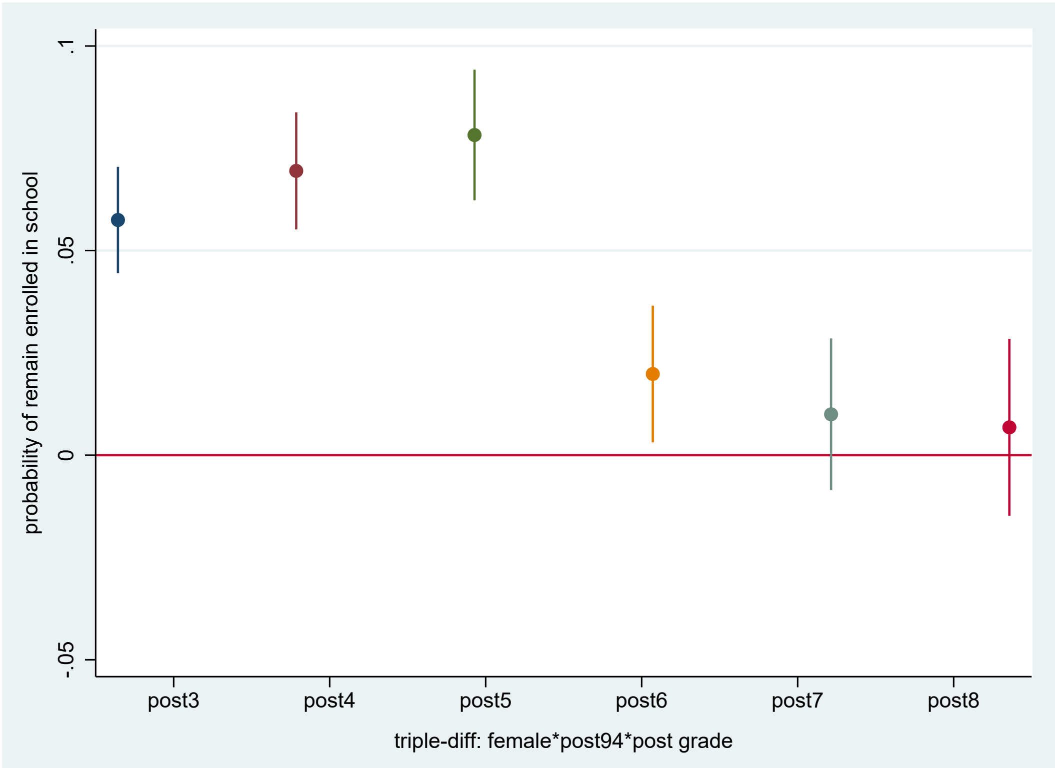


Figure 5: The replication of coefficient of triple-difference term with the household fixed effects presented in Table 9 in Heath and Mobarak (2015) by redefining the cutoff grade using 2010 HIES data.

between boys and girls and between pre- and post-1994 periods. On the contrary, Heath and Mobarak (2015) take the difference between those who had at least 6 years of education ("post6") and those who did not in a given year. Therefore, the child in their classification is already enrolled in secondary school and the potential impact of FSPs on the enrollment decision into grade 6 among the primary graduates is not taken into account. We argue that our choice would be more appropriate because the FSPs affect the enrollment decision as long as the child has graduated from primary schools. That is, the FSPs makes it more attractive to keep girls enrolled in school after the completion of primary education. We demonstrate this point by altering the cutoff completed grade, which is essentially taken as the eligibility for the FSPs, for the third difference in Figure 5 using HIES 2010 survey data and household fixed effects are controlled in the analysis.

While Figure 5 shows positive and significant effect at "post5", it is not immediately clear how this result can be compared with Heath and Mobarak (2015). We, therefore, repeated the same analysis with the data used by Heath and Mobarak (2015). As Figure 6 shows, all the point estimates are insignificant. But more importantly, the figure is broadly consistent with Figure 5 and indicates that their finding of a negative point estimate is driven by their choice of cutoff grade at grade 6. We also repeated our analysis with a subsample from Dhaka and Gazipur districts to make the data we use more comparable with those used by Heath and Mobarak (2015). As Figure 7 shows, we again find a very

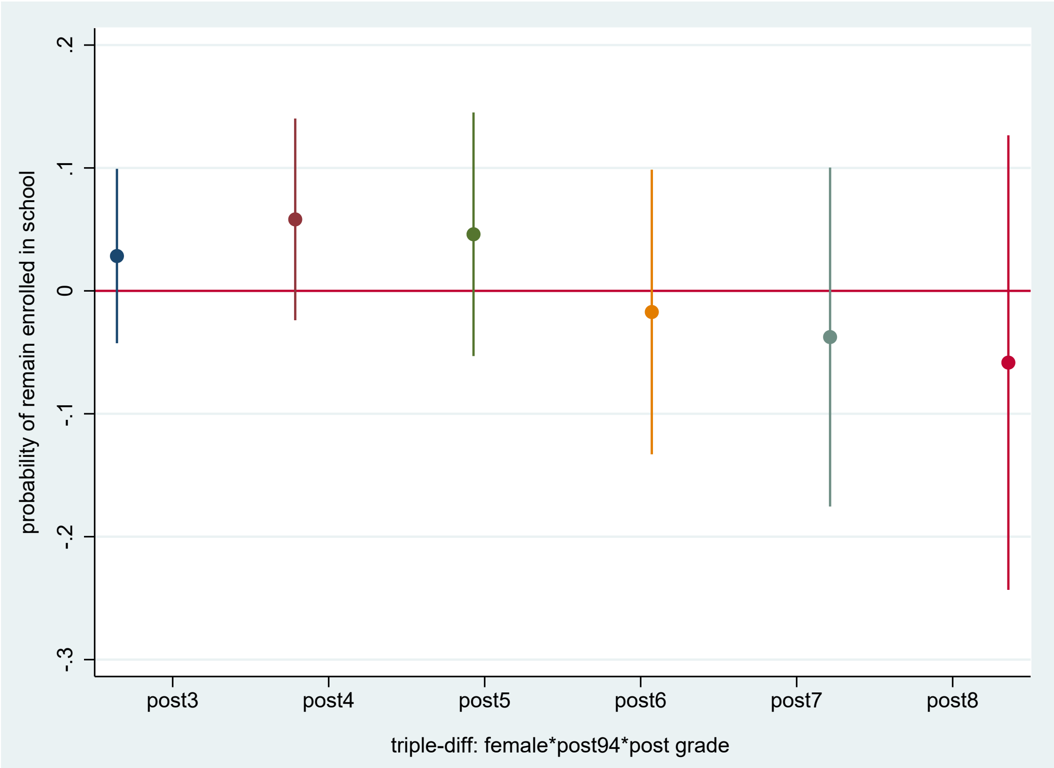


Figure 6: The replication of coefficient of triple-difference term with household fixed effect presented in Table 9 in Heath and Mobarak (2015) by redefining the cutoff grade using the data used in Heath and Mobarak (2015).

similar pattern. It also shows that the impact of FSPs is significantly positive when the completion of grade 5 is used as a cutoff.

The pattern we observe is indeed to be expected. Because the decision of the grade-5 completers to enroll in a secondary school is affected by the FSPs, the apparent impact of FSPs would go down if they are treated as FSP-ineligible individuals as Heath and Mobarak (2015). Because there are requirements for continuation and because it is in general difficult to go back to school once the child drops out of school, we expect that the impact of FSPs to diminish for higher grades. The findings in Figures 5-7 indeed support this argument.

# E Additional tables for summary statistics and detailed regression results

Table 15 and 16 provide the same summary statistics as Tables 2 and 3 except that the former are for the years 2000 and 2005. In Table 17, we provide the complete regression results for the three part model presented in Table 4. The estimated values of $\rho$'s were all highly statistically significant, indicating the relevance of allowing for the correlation in the error terms. The estimations for $\rho_{dy}$ and

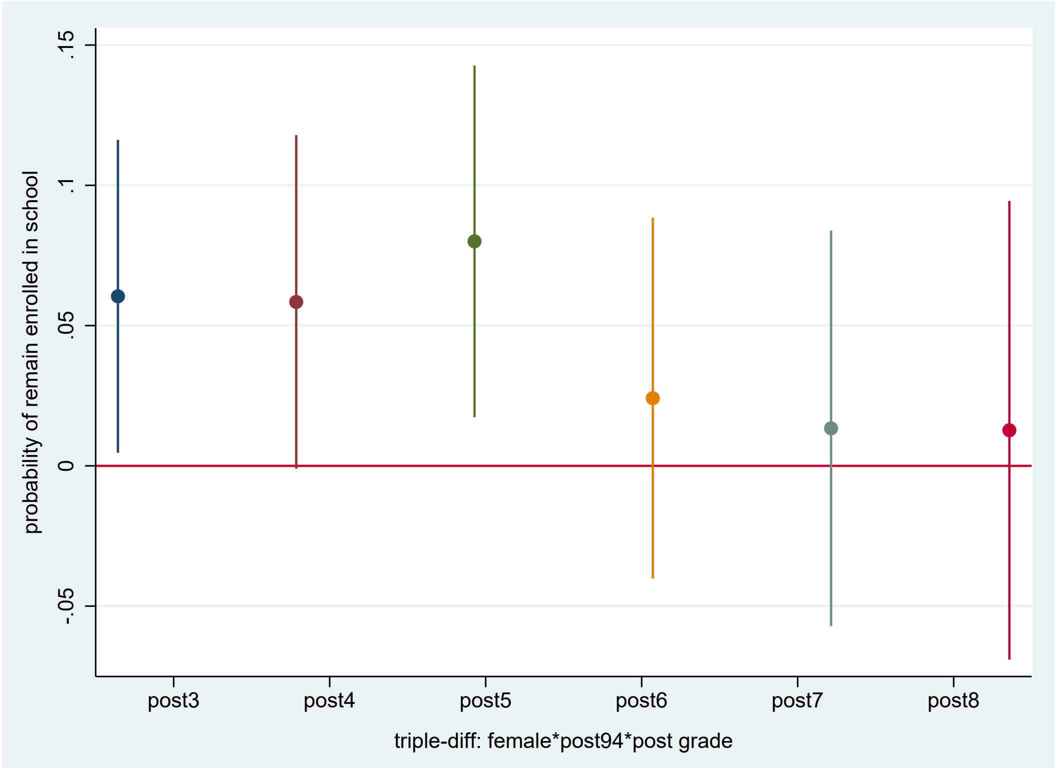


Figure 7: The replication of coefficient of triple-difference term with household fixed effect presented in Table 9 in Heath and Mobarak (2015) by redefining the cutoff grade using 2010 HIES data for Dhaka and Gazipur districts. Heath and Mobarak (2015) uses 2009 survey data in sub-districts of Dhaka and Gazipur districts.

$\rho_{ds}$ are negatively significant at a 1 percent level from 2000 onwards. One plausible explanation is that the unobserved academic capability affects the enrollment and the other two decisions in different directions, possibly because very smart students need little spending on education. This possibility appears to be consistent with our estimate of $\rho_{ys}$, which is positive and significant from 2000 onwards. Table 18 show that the regression results are similar when the independence of error terms is assumed. The sign and significance of the coefficients remain similar but the absolute value of the coefficient for the conditional education expenditure and core share equations appears to be larger when the dependence structure is allowed for. We then present the marginal effects of the girl dummy when the independence of error terms is assumed in Table 19. The results are qualitatively similar to Table 9. Finally, Table 20 provides the marginal effects of the girl and FSP dummies at the sample mean for each education expenditure item.

Table 15: Summary statistics of basic covariates by gender for 2000 and 2005 secondary-school age group

| | 2000 | | | | 2005 | | | |
|---|---|---|---|---|---|---|---|---|
| | **Boy (B)** | **Girl (G)** | **Diff (G-B)** | **All** | **Boy (B)** | **Girl (G)** | **Diff (G-B)** | **All** |
| Variables | (1) | (2) | (2)-(1) | (4) | (5) | (6) | (6)-(5) | (8) |
| ***All children aged 11-15*** | | | | | | | | |
| Enrolled in secondary school | 0.331 | 0.444 | 0.113 | 0.386 | 0.407 | 0.509 | 0.102 | 0.457 |
| | (0.471) | (0.497) | *** | (0.487) | (0.491) | (0.500) | *** | (0.498) |
| Child's age (yrs) | 13.012 | 12.908 | -0.104 | 12.961 | 13.079 | 13.001 | -0.078 | 13.041 |
| | (1.401) | (1.342) | *** | (1.373) | (1.400) | (1.350) | ** | (1.376) |
| HH per capita expenditure | 10.722 | 11.419 | 0.697 | 11.064 | 14.296 | 14.717 | 0.421 | 14.504 |
| | (7.809) | (9.013) | *** | (8.428) | (10.282) | (11.578) | | (10.943) |
| Household size | 6.405 | 6.559 | 0.154 | 6.480 | 5.990 | 6.102 | 0.112 | 6.046 |
| | (2.347) | (2.392) | ** | (2.371) | (2.232) | (2.162) | * | (2.198) |
| Father's education (yrs) | 2.841 | 3.104 | 0.263 | 2.970 | 3.045 | 3.186 | 0.141 | 3.115 |
| | (4.130) | (4.198) | ** | (4.165) | (4.187) | (4.208) | | (4.198) |
| Mother's education (yrs) | 1.725 | 1.939 | 0.214 | 1.830 | 2.224 | 2.322 | 0.098 | 2.272 |
| | (3.095) | (3.198) | ** | (3.148) | (3.501) | (3.532) | | (3.517) |
| Number of children | 3.533 | 3.635 | 0.102 | 3.583 | 3.243 | 3.336 | 0.093 | 3.289 |
| | (1.741) | (1.758) | ** | (1.750) | (1.568) | (1.584) | ** | (1.576) |
| Urban | 0.318 | 0.339 | 0.021 | 0.328 | 0.341 | 0.342 | 0.001 | 0.341 |
| | (0.466) | (0.473) | | (0.470) | (0.474) | (0.474) | | (0.474) |
| Female head | 0.073 | 0.080 | 0.007 | 0.076 | 0.095 | 0.093 | -0.002 | 0.094 |
| | (0.260) | (0.271) | | (0.265) | (0.293) | (0.290) | | (0.292) |
| Head is a wage worker | 0.381 | 0.393 | 0.012 | 0.387 | 0.414 | 0.444 | 0.030 | 0.429 |
| | (0.486) | (0.488) | | (0.487) | (0.493) | (0.497) | ** | (0.495) |
| Head's age (yrs) | 46.988 | 46.877 | -0.111 | 46.933 | 47.671 | 47.602 | -0.069 | 47.637 |
| | (10.738) | (10.957) | | (10.845) | (10.623) | (10.445) | | (10.535) |
| Muslim | 0.919 | 0.922 | 0.003 | 0.921 | 0.890 | 0.893 | 0.003 | 0.892 |
| | (0.272) | (0.268) | | (0.270) | (0.313) | (0.309) | | (0.311) |
| Hindu | 0.076 | 0.071 | -0.005 | 0.073 | 0.093 | 0.093 | 0.000 | 0.093 |
| | (0.265) | (0.256) | | (0.261) | (0.290) | (0.290) | | (0.290) |
| *Obs* | 2,488 | 2,390 | | 4,878 | 2,848 | 2,790 | | 5,638 |
| ***Enrolled in secondary school children aged 11-15*** | | | | | | | | |
| Govt school | 0.25 | 0.23 | -0.02 | 0.24 | 0.25 | 0.23 | -0.02 | 0.24 |
| | (0.44) | (0.42) | | (0.43) | (0.44) | (0.42) | | (0.43) |
| Private school | 0.68 | 0.70 | 0.02 | 0.69 | 0.66 | 0.69 | 0.03 | 0.67 |
| | (0.47) | (0.46) | | (0.46) | (0.47) | (0.46) | | (0.47) |
| Other | 0.07 | 0.07 | 0.00 | 0.07 | 0.09 | 0.09 | 0.00 | 0.09 |
| | (0.25) | (0.25) | | (0.25) | (0.28) | (0.28) | | (0.28) |
| *Obs* | 824 | 1,061 | | 1,885 | 1,159 | 1,420 | | 2,579 |

Note: Standard errors are reported in parentheses below the mean. ***,** ,* denote that the means of girl and boy are different at 1, 5, 10 percent significant level, respectively. The unit for household per capita expenditure is thousand taka. Other in school type include other types of schools, such as religious schools (like madrasa) and NGO schools. The summary statistics of basic covariates for 2000 and 2005 are presented in Table 2 in Section 4.

Table 16: Summary statistics of education expenditure by items for secondary school enrollees in 2000 and 2005

| | 2000 | | | | 2005 | | | |
|---|---|---|---|---|---|---|---|---|
| | **Boy (B)** | **Girl (G)** | **Diff (G-B)** | **% Zeros** | **Boy (B)** | **Girl (G)** | **Diff (G-B)** | **% Zeros** |
| Taka | (1) | (2) | (2)-(1) | (4) | (5) | (6) | (6)-(5) | (8) |
| **Core** | 2,116.3 | 1,681.4 | -434.9 | 1% | 2,785.7 | 2,377.5 | -408.2 | 1% |
| | (2,091.7) | (1,723.0) | *** | | (2,451.5) | (2,405.2) | *** | |
| *Tuition* | 320.7 | 131.0 | -189.7 | 48% | 374.2 | 162.2 | -212.0 | 50% |
| | (384.0) | (308.0) | *** | | (489.2) | (491.9) | *** | |
| *Home Tutor* | 1,031.2 | 821.2 | -210 | 49% | 1,437.7 | 1,289.4 | -148.3 | 42% |
| | (1,665.1) | (1,366.7) | *** | | (2,056.6) | (2,046.3) | * | |
| *Material* | 764.4 | 729.2 | -35.2 | 1% | 973.8 | 925.9 | -47.9 | 1% |
| | (526.6) | (486.2) | | | (589.1) | (579.0) | ** | |
| **Peripheral** | 917.5 | 877.9 | -39.6 | 1% | 1,166.4 | 1,067.5 | -98.9 | 0% |
| | (930.0) | (901.0) | | | (1,383.4) | (1,103.1) | ** | |
| *Admission* | 170.1 | 152.4 | -17.7 | 26% | 202.4 | 188.8 | -13.6 | 26% |
| | (232.6) | (218.4) | * | | (310.1) | (349.0) | | |
| *Exam* | 152.1 | 142.5 | -9.6 | 4% | 173.0 | 177.6 | 4.6 | 4% |
| | (166.4) | (121.1) | | | (137.1) | (180.8) | | |
| *Uniform* | 238.7 | 257.2 | 18.5 | 46% | 342.8 | 343.7 | 0.9 | 35% |
| | (315.2) | (291.8) | | | (450.0) | (390.6) | | |
| *Meal* | 175.9 | 176.3 | 0.4 | 63% | 193.2 | 155.3 | -37.9 | 68% |
| | (368.1) | (349.1) | | | (408.5) | (359.0) | ** | |
| *Transportation* | 120.8 | 111.3 | -9.5 | 84% | 119.1 | 129.1 | 10.0 | 86% |
| | (420.1) | (401.1) | | | (491.6) | (506.5) | | |
| *Others* | 59.8 | 38.2 | -21.6 | 75% | 136.0 | 73.1 | -62.9 | 66% |
| | (312.3) | (214.4) | * | | (793.5) | (251.0) | *** | |
| **Total** | 3,033.8 | 2,559.3 | -474.5 | | 3,952.1 | 3,445.0 | -507.1 | |
| | (2,664.8) | (2,318.8) | *** | | (3,127.3) | (2,979.2) | *** | |
| **Core Share** | 0.68 | 0.63 | -0.05 | | 0.69 | 0.65 | -0.04 | |
| | (0.18) | (0.20) | *** | | (0.18) | (0.19) | *** | |
| Obs | 824 | 1,061 | | | 1,159 | 1,420 | | |

Note: Standard errors are reported in parentheses below the mean. ***,** ,* denote that the means of girl and boy are different at 1, 5, 10 percent significant level, respectively. The summary statistics is for subsample of children who were enrolled in secondary school at the time of survey. Core share stands for the ratio of core components over total education expenditure. The admission fee here include admission, annual session and registration fee based on the component classification in 1995 HES questionnaire. The summary statistics of education expenditure by items for enrolled secondary-school age group in 1995 and 2010 are presented in Table 3 in Section 4.

Table 17: ML Estimation of Three-part Model with Dependence for Secondary-school Age Group

| Secondary Age (11-15) | **1995** | | | **2000** | | | **2005** | | | **2010** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Coef.* | $d$ | Cond $y$ | Cond $s$ | $d$ | Cond $y$ | Cond $s$ | $d$ | Cond $y$ | Cond $s$ | $d$ | Cond $y$ | Cond $s$ |
| Girl | -0.001 | -0.085*** | 0.001 | 0.339*** | -0.174*** | -0.082*** | 0.291*** | -0.154*** | -0.071*** | 0.289*** | -0.131*** | -0.067*** |
| | (0.042) | (0.032) | (0.032) | (0.039) | (0.049) | (0.014) | (0.034) | (0.027) | (0.012) | (0.033) | (0.025) | (0.009) |
| Log(per capita exp) | 0.505*** | 0.755*** | -0.326 | 0.480*** | 0.793*** | -0.124*** | 0.374*** | 0.609*** | -0.078*** | 0.357*** | 0.701*** | -0.046** |
| | (0.047) | (0.043) | (0.261) | (0.050) | (0.052) | (0.041) | (0.044) | (0.034) | (0.027) | (0.046) | (0.034) | (0.022) |
| Log(hh size) | 0.090 | 0.116* | -0.033 | 0.142* | 0.222*** | -0.014 | -0.089 | 0.139*** | 0.025 | 0.124 | 0.317*** | -0.049** |
| | (0.083) | (0.068) | (0.052) | (0.084) | (0.068) | (0.025) | (0.075) | (0.053) | (0.019) | (0.082) | (0.064) | (0.021) |
| Father edu (yrs) | 0.081*** | 0.014** | -0.006 | 0.073*** | -0.005 | -0.008*** | 0.062*** | 0.006 | -0.009*** | 0.039*** | 0.010** | -0.004*** |
| | (0.007) | (0.006) | (0.006) | (0.007) | (0.011) | (0.002) | (0.006) | (0.005) | (0.002) | (0.006) | (0.004) | (0.001) |
| Mother edu (yrs) | 0.088*** | 0.025*** | -0.010 | 0.068*** | -0.007 | -0.008*** | 0.066*** | 0.023*** | -0.009*** | 0.073*** | 0.017*** | -0.008*** |
| | (0.010) | (0.007) | (0.009) | (0.009) | (0.010) | (0.002) | (0.008) | (0.006) | (0.002) | (0.007) | (0.005) | (0.002) |
| No. of children | 0.008 | -0.004 | -0.001 | -0.026* | -0.022 | 0.003 | -0.015 | -0.024* | -0.006 | -0.028 | -0.021 | 0.009** |
| | (0.016) | (0.013) | (0.007) | (0.015) | (0.014) | (0.004) | (0.017) | (0.013) | (0.004) | (0.018) | (0.016) | (0.004) |
| Urban | -0.082 | 0.257*** | -0.097 | -0.121** | 0.287*** | 0.000 | -0.095** | 0.275*** | -0.011 | -0.140*** | 0.233*** | 0.052*** |
| | (0.053) | (0.045) | (0.091) | (0.048) | (0.046) | (0.020) | (0.040) | (0.030) | (0.016) | (0.040) | (0.029) | (0.012) |
| Female head | -0.038 | -0.069 | -0.019 | 0.106 | 0.004 | -0.027 | 0.030 | 0.097* | 0.016 | -0.062 | 0.103** | -0.004 |
| | (0.084) | (0.074) | (0.044) | (0.078) | (0.075) | (0.022) | (0.068) | (0.050) | (0.019) | (0.058) | (0.046) | (0.014) |
| Head is a wage worker | -0.103** | 0.029 | -0.044* | -0.201*** | 0.152*** | 0.028* | -0.210*** | 0.086*** | 0.037*** | -0.145*** | 0.023 | 0.015* |
| | (0.049) | (0.041) | (0.023) | (0.045) | (0.052) | (0.014) | (0.038) | (0.030) | (0.011) | (0.038) | (0.029) | (0.009) |
| Head's age (yrs) | 0.000 | -0.003* | 0.001 | 0.001 | -0.005*** | -0.001 | -0.004** | 0.001 | -0.000 | -0.004** | -0.003** | 0.000 |
| | (0.002) | (0.002) | (0.001) | (0.002) | (0.002) | (0.001) | (0.002) | (0.001) | (0.001) | (0.002) | (0.001) | (0.000) |
| Muslim | -0.239 | -0.138 | 0.079 | 0.219 | 0.128 | 0.143** | -0.073 | -0.246*** | 0.000 | 0.199 | -0.219* | 0.039 |
| | (0.244) | (0.138) | (0.075) | (0.312) | (0.186) | (0.064) | (0.167) | (0.094) | (0.045) | (0.211) | (0.123) | (0.052) |
| Hindu | -0.062 | -0.128 | 0.086 | 0.306 | 0.263 | 0.140** | -0.132 | -0.161 | 0.020 | 0.216 | -0.189 | 0.070 |
| | (0.253) | (0.144) | (0.075) | (0.318) | (0.197) | (0.067) | (0.177) | (0.101) | (0.046) | (0.217) | (0.128) | (0.053) |
| Secondary school accessibility | 2.454*** | | | 5.940*** | | | 1.487*** | | | 2.501*** | | |
| | (0.606) | | | (1.019) | | | (0.364) | | | (0.447) | | |
| Madrasa school accessibility | -0.287 | | | -6.142*** | | | 0.274 | | | 0.763 | | |
| | (0.912) | | | (1.045) | | | (0.463) | | | (0.535) | | |
| Government school | | 0.160 | | | 0.208** | | | 0.135** | | | 0.275*** | |
| | | (0.121) | | | (0.092) | | | (0.058) | | | (0.058) | |
| Private school | | 0.195 | | | 0.387*** | | | 0.286*** | | | 0.430*** | |
| | | (0.128) | | | (0.083) | | | (0.054) | | | (0.052) | |
| Log(education expend) | | | 0.449 | | | 0.068 | | | 0.061 | | | 0.032 |
| | | | (0.343) | | | (0.050) | | | (0.043) | | | (0.029) |
| $\sigma_y$ | | 0.681*** | | | 0.740*** | | | 0.650*** | | | 0.671*** | |
| | | (0.017) | | | (0.056) | | | (0.017) | | | (0.015) | |
| $\sigma_s$ | | 0.321* | | | 0.242*** | | | 0.257*** | | | 0.235*** | |
| | | (0.190) | | | (0.013) | | | (0.010) | | | (0.010) | |
| $\rho_{dy}$ | | 0.192** | | | -0.456** | | | -0.196*** | | | -0.285*** | |
| | | (0.076) | | | (0.192) | | | (0.072) | | | (0.075) | |
| $\rho_{ds}$ | | -0.165 | | | -0.829*** | | | -0.935*** | | | -0.894*** | |
| | | (0.116) | | | (0.081) | | | (0.025) | | | (0.029) | |
| $\rho_{ys}$ | | -0.810*** | | | 0.146 | | | 0.100 | | | 0.221** | |
| | | (0.250) | | | (0.178) | | | (0.124) | | | (0.094) | |
| Observations | | 5,011 | | | 4,878 | | | 5,638 | | | 6,205 | |

Note: ***,** ,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at household level are reported in parentheses. The estimations are obtained using three-part model constructed in Section 3. Secondary school (madrasa) accessibility is the number of secondary schools (madrasa) per 1000 people by population size. It is calculated at district level for all years except 2000, which is at subdivision level. Additional controls include age dummies. $\sigma_y$ and $\sigma_s$ are the standard errors of conditional expenditure and core share, respectively. $\rho_{dy}$ is the correlation between enrollment and conditional expenditure. Similar to $\rho_{ys}$ and $\rho_{ds}$. Detailed results for Table 4 in Section 5.

Table 18: ML Estimation of Three-part Model of Different Error Structure

| | $d$ | Cond $y$ | Cond $s$ |
|---|---|---|---|
| **1995** | | | |
| Independence | -0.003 | -0.086*** | -0.030*** |
| | (0.042) | (0.032) | (0.009) |
| Dependence | -0.001 | -0.085*** | 0.001 |
| | (0.042) | (0.032) | (0.032) |
| **2000** | | | |
| Independence | 0.331*** | -0.111*** | -0.047*** |
| | (0.041) | (0.032) | (0.009) |
| Dependence | 0.339*** | -0.174*** | -0.082*** |
| | (0.039) | (0.049) | (0.014) |
| **2005** | | | |
| Independence | 0.309*** | -0.131*** | -0.027*** |
| | (0.037) | (0.025) | (0.007) |
| Dependence | 0.291*** | -0.154*** | -0.071*** |
| | (0.034) | (0.027) | (0.012) |
| **2010** | | | |
| Independence | 0.295*** | -0.101*** | -0.031*** |
| | (0.035) | (0.024) | (0.006) |
| Dependence | 0.289*** | -0.131*** | -0.067*** |
| | (0.033) | (0.025) | (0.009) |

Note: $^{***}$,$^{**}$,$^{*}$ denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at household level are reported in parentheses. The estimations are obtained using three-part model constructed in Section 3. The independence model assumes $\rho_{dy}, \rho_{ds}$ and $\rho_{ys}$ are all zeros. Additional covariates are the same as discussed in Table 4.

Table 19: Marginal Effects of the Girl Dummy at the Mean

| | *Marginal effects at the mean* | E($d$) (1) | E($y \mid d=1$) (2) | E($y$) (3) | E($ys \mid d=1$) (4) | E($ys$) (5) |
|---|---|---|---|---|---|---|
| | Girl | -0.001 | -179.245*** | -40.798 | -178.410*** | -27.547*** |
| **1995** | | (0.015) | (67.804) | (25.872) | (48.393) | (10.206) |
| | *Obs* | 5,011 | 1,798 | 5,011 | 1,798 | 5,011 |
| | Girl | 0.124*** | -269.487*** | 134.390*** | -285.579*** | 34.790** |
| **2000** | | (0.015) | (79.523) | (31.346) | (55.136) | (16.395) |
| | *Obs* | 4,878 | 1,885 | 4,878 | 1,885 | 4,878 |
| | Girl | 0.122*** | -433.665*** | 154.744*** | -374.411*** | 32.632 |
| **2005** | | (0.014) | (84.342) | (45.672) | (60.282) | (20.393) |
| | *Obs* | 5,638 | 2,579 | 5,638 | 2,579 | 5,638 |
| | Girl | 0.117*** | -612.691*** | 304.818*** | -582.444*** | 44.740 |
| **2010** | | (0.014) | (145.621) | (88.339) | (103.683) | (38.249) |
| | *Obs* | 6,205 | 3,172 | 6,205 | 3,172 | 6,205 |

Note: ***,**,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at household level are reported in parentheses. The marginal effects are calculated using marginal effect code in STATA based on three-part model with independence estimates.E(·) stands for the expectation of the variable in the brackets. Estimates in column (1) are the marginal effect of the girl dummy on the expected enrollment in secondary school for the children in the secondary-school age group. The marginal effects presented in Columns (2) to (5) are in taka in nominal terms. Unconditional [Conditional] expectations are evaluated at the mean of the full sample [subsample of secondary-school enrollees]. The marginal effects at the sample mean for three-part model constructed in Section 3 are presented in Table 9 in Section 5.

Table 20: Tobit Regressions of Education Expenditure Items for Secondary School Enrollees with FSP Dummy

| | Marginal effects at the mean | **Core** (1) | Tuition (2) | Home Tutor (3) | Material (4) | **Peripheral** (5) | Admission (6) | Exam (7) | Uniform (8) | Meal (9) | Transport (10) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | FSP | -72.9 | -518.2*** | 226.1 | 101.4*** | 125.5** | -69.5*** | 14.8* | 95.2*** | 136.1** | 283.4** |
| | | (95.5) | (52.3) | (153.4) | (29.2) | (51.2) | (18.5) | (7.8) | (31.4) | (53.0) | (137.6) |
| **2000 Sec** | Girl | -240.5** | -207.7*** | -337.4** | -66.0** | -44.1 | 20.9 | -11.2 | 29.5 | -39.1 | -182.8 |
| | | (94.5) | (35.9) | (141.3) | (27.1) | (45.5) | (18.8) | (7.3) | (27.8) | (48.8) | (126.5) |
| | *Obs* | 1885 | 1885 | 1885 | 1885 | 1885 | 1885 | 1885 | 1885 | 1885 | 1885 |
| | FSP | -112.3 | -704.9*** | 245.3 | 64.8** | 4.7 | -70.2*** | 1.7 | 61.3** | 67.8 | 163.7 |
| | | (107.7) | (86.6) | (154.7) | (28.0) | (54.3) | (23.4) | (9.8) | (30.8) | (55.5) | (151.2) |
| **2005 Sec** | Girl | -202.1** | -383.7*** | -228.7 | -56.4** | -47.4 | 20.7 | 8.7 | -6.4 | -87.9* | -26.8 |
| | | (99.4) | (51.0) | (139.7) | (25.1) | (58.5) | (22.5) | (9.3) | (29.5) | (51.6) | (135.5) |
| | *Obs* | 2579 | 2579 | 2579 | 2579 | 2579 | 2579 | 2579 | 2579 | 2579 | 2579 |

Note: ***,**,* denote statistical significance at 1, 5, 10 percent levels. Standard errors clustered at household level are reported in parentheses. Marginal effects at mean using Tobit regressions of education expenditure items for subsample of enrolled children are reported. The covariates are the same to those used in column Cond $y$ of Table 4. The admission fee here include admission, annual session and registration fee to be consistent with Table 3.