This is the accepted version of the following article: Himaz, R and H. Aturupane (forthcoming), Why are boys falling behind? Explaining gender gaps in school attainment in Sri Lanka, *World Development, 2021*

Why are boys falling behind? Explaining gender gaps in school attainment in Sri Lanka

Rozana Himaz* and Harsha Aturupane**

**** Business School, Oxford Brookes University, UK ** World Bank, Washington D.C.***

Abstract

An increasingly common trend in developed countries and middle income countries such as Thailand, South Africa, Malaysia, Indonesia and Sri Lanka is that females outperform males in terms of attainment at school and enrollment in higher education, on average. Surprisingly, in countries such as Sri Lanka and Thailand, households also seem to allocate significantly higher resources towards girls' education rather than boys' (Himaz, 2010; Wongmonta & Glewwe, 2017). This paper looks at attainment in mathematics among a sample of 12 year olds in Sri Lanka and uses decomposition analysis to see to what extent parental aspirations and teacher attitudes, *inter alia*, can explain gender differentials disfavoring boys. The paper finds that although teacher attitudes and parental aspirations are significantly lower for boys, these factors -as we measure them- do not sufficiently explain the attainment gap. Much of the gap remains 'unexplained' and is due to differences in *returns* to endowments. The paper argues that positive discrimination of men in the labor market and bottle necks in higher education may be important in understanding the unexplained component. This emphasizes the need to look beyond school level interventions to address the issue of gender imbalances in attainment.

Keywords: Sri Lanka, South Asia, Gender, Education, attainment, bias.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

We would like to thank two anonymous referees and the editor for very helpful comments and suggestions. We are also grateful to seminar participants at Verite Research (Asia) for their useful insights and discussion. All potential errors are our own. This version replaces any working paper versions of the paper.

Highlights

- Male attainment at school and enrolment in post-compulsory schooling is significantly below those of females in Sri Lanka
- Implicit biases favour girls with parental aspirations and teacher attitudes higher for girls than boys
- Gaps in maths scores at age 12 are explained by differences in returns to endowments rather than differences in endowments
- Earnings foregone are significant in a male's decision to enrol in school even when controlled for wealth.
- Positive discrimination of males in the labour market may explain poor attainment at the school level

1. Introduction

An emerging trend in middle income countries such as Malaysia, Thailand, Sri Lanka, the Philippines, South Africa and Burma, matching similar trends in high income OECD countries, is the significant lagging behind of men compared with women, on average, in terms of enrollment in post-compulsory schooling, participation in tertiary education (see Appendix Figures A1-A2) and attainment at school. Disparities in Sri Lanka in this regard are particularly high. For example, 63 per cent of all state university enrollments in 2015 comprised females, across all disciplines including medicine, science, arts and management with the only exception being engineering where the female to male enrollment ratio was 35:65. Female participation was also high at other tertiary institutions with the female to male ratio being 61:39[1]. This divergence is clear at almost all levels of education in various dimensions. Senior secondary school completion rates (i.e., complete grade 13), for instance, are higher for females than males and correspondingly the dropout rate for boys are higher at 33 per cent compared to 28 per cent in girls. Moreover, attainment in grade 8 while at junior secondary is lower for boys rather than girls based on standardized test scores for Mathematics, Science and English (Dundar et al., 2017)[2]. Even at the primary school level, girls seem to outperform boys in terms of attainment in mathematics and language (Sinhala or Tamil) with entrance scores based on the national grade five scholarship exams for the most prestigious girls’ schools being higher than for boys' schools in the past few years. Quite notably (and perhaps surprisingly) in countries such as Sri

[1] Own calculations based on University Grant Commission 2015 data.

[2] This paper uses the following terminology: Primary refers school to grades 1-5, secondary to school grades 6-13 (subdivided as junior secondary grades 6-9 and senior secondary grades 10-13). Students in Grade 1 are around age 5/6 and grade 13 around 18/19 and even 20 if they repeat the grade.

Lanka and Thailand, households seem to allocate significantly higher resources towards girls' education rather than boys especially in the run up to major public exams (Himaz, 2010; Wongmonta & Glewwe, 2017). This may be contributing to the higher performance of girls whilst also being reflective of attitudes within households towards male versus female engagement in formal education.

There are various explanations offered in the literature, particularly those pertaining to OECD countries, as to why boys may be falling behind. The views draw from various subjects including neurology, psychology, pedagogy, economics, feminism, politics, history, geography and anthropology. They can be roughly grouped as pertaining to individual, societal/labor market-oriented and school environment-related aspects. At the individual level, a key contributor to differential attainment between boys and girls may be due to them facing different cognitive and non-cognitive developmental timetables, on average, during infancy due to neurological and hormonal factors (Schore, 2017). Indeed by most measures of sensory and cognitive development girls are more advanced than boys during infancy with vision, hearing, memory, smell and touch all being more acute for girls than boys. Thus by the time they start school, girls tend to exhibit more 'readiness' for school with more developed language and non-cognitive skills compared to boys. These differences in early childhood may influence parental, teacher and even child attitudes setting boys up on a trajectory of lower attainment (Cornwell, Mustard, & Van Parys, 2013). Although environmental factors can accentuate or reduce differences in the attainment of skills, implicit biases can affect child outcomes even at the preschool stage (Gilliam, Maupin, Reyes, Accavitti, & Shic, 2016).

Contributing to differences in attitudes are societal factors and gender stereotyping. For example, boys maybe subject to lower parental supervision (Goolamally and Ahmad

(2010) from (UNGEI, 2012)) and lower teacher expectations regarding ‘responsible’ behavior (Abdullah, 2009). Thus socially acceptable male behaviors may often be considered incongruent with academic achievement. Participation in education beyond compulsory years can also be lower for men if there is a higher expectation for men to contribute to household incomes rather than women.

Another less-researched explanation, particularly applicable to Sri Lanka and possibly other developing countries, is supply-side bottle necks in higher education that can also influence attitudes to education at a younger age. For example, if university places are very competitive and limited then investing time and effort to secure a place in higher education maybe seen as less preferable to seeking employment soon after compulsory education. Significant gender based variations in such attitudes to learning can occur if alternative opportunities offered in the labor market vary by gender. Thus if men face a wider range of employment opportunities, are paid more for the same set of skills or have more opportunities to rent-seek through joining political organizations, it can all manifest in lower educational outcomes, poorer attitudes to education and less incentive to participate in higher education on average among men compared to women.

Other explanations focus on classroom-level factors including learning styles, teaching styles, curriculum, teacher and student absenteeism as well as fewer male teachers (and therefore fewer male role models in education). There could also be more systemic factors such as how education is delivered. For example, centralized education systems may give schools less power and control to make decisions that suit their particular context, compared to decentralized education systems that provide a mechanism for collaboration between schools and local communities. School based management programmes, for instance, that transfer decision making authority of some aspects of school operations to local community members and stakeholders can support the improvement of boys' performance especially if their falling behind is recognized to be a concern in the local context[3].

There is little research specific to developing countries regarding the issue of gender gaps disfavoring boys apart from reports such as UNICEF (2006) that raise the issue and suggest free education and strategies to improve school attendance as ways forward, *inter alia*. The case of Sri Lanka offers a good opportunity to glean some insights into this issue. Education has been 'free' since 1948. Therefore some of the financial cost related issues pertaining to accessing education are lower in Sri Lanka compared to other developing countries. Moreover high enrollment and gender parity in school enrollment were targets achieved many decades ago. In spite of such strides, gender inequalities disfavoring boys have emerged and continue to rise. The Sri Lankan experience, therefore, can offer some useful insights to other developing countries in terms of what may matter the most when implementing gender sensitive policy.

[3] Most of the insights as to how to improve the performance of boys also come from the literature pertaining to OECD countries. They focus mainly on influencing attitudes to reduce negative gender stereotyping within schools and communities. Thus specific strategies to improve boys performance include the use of peer groups, family and community to help smooth transitions over the school trajectory to reduce bullying and the setting in of negative stereotyping, early intervention programs through community based parenting efforts and the involvement of fathers/males and other school-based interventions that include attracting more male teachers (Hillman & Robinson, 2016; Younger, Warrington, & McLellan, 2005). There is also the recognition that focusing simply on education sector reforms is insufficient for what is a complex issue and policy efforts should be collaborative, encouraging coherence across policies including economic, labor, gender rights and finance policies (UNESCO, 2015).

The issue is complex so this paper uses two approaches to identify some of the underlying causes in the Sri Lankan context. It considers first whether specific household characteristics such as parental aspirations, teacher characteristics such as gender and attitudes, and school characteristics such as whether the school based management initiative (the Programme for School Improvement, PSI) is implemented, can explain maths scores among boys and girls in grade seven (aged 12) in Sri Lanka. To understand the extent to which differences in endowments versus returns to endowments can explain gaps in attainment we decompose results using Oaxaca-Blinder decomposition analysis. To check the robustness of results the paper also looks at the impact of the various characteristics across the distribution of mathematics scores by decomposing using recentered influence function regressions (S. P. Firpo, Fortin, & Lemieux, 2018). The data come from a short panel that follows the same schools (rather than the same individuals) in 2012 and 2016, collected by the World Bank. The second approach is to go beyond factors limited to the compulsory the school system and look at post-compulsory education and the labour market. Thus the paper uses data from the 2012/13 Household Income and Expenditure Survey collected by the Department of Census and Statistics to see what drives lower male enrolment beyond the compulsory years of schooling (age 14). This analysis also looks at gender-based asymmetries in the youth labour market and the higher education system. We also look briefly at the role dowries at marriage may play in education investment. The results from the first and second approaches are then pieced together to draw insights as to what may explain why boys fall behind in the case of a developing country such as Sri Lanka.

The paper is organized as follows. The next section looks at the data used and provides descriptive statistics. Section 3 presents the empirical specification for attainment in mathematics at age 12 and the decomposition analysis while section 4 discusses the results. Section 5 looks at enrollment beyond compulsory schooling, labor market asymmetries and selected issues pertaining to the marriage market. Section 6 concludes.

2. Data and Descriptive Stats

The data used for looking at what explains gender-based gaps in math scores amongst 12 year olds come from a longitudinal survey that followed the same schools over time, collecting data in each school for a class of grade 7 students, their households, and teachers and school-level characteristics. The original purpose of the two rounds of data collection was to assess the effectiveness of the PSI implemented by the Government. The rich dataset, however, makes it very suitable to explain what may drive math scores and gender-based differences therein[4].

The first round of data was collected in 2012 from 72 schools across all 9 provinces of the country. The schools represented a mix of the various school types (1AB, IC or 2 based on curriculum delivered)[5], levels (based on facilities available)[6] and races (Tamil, Sinhala, Muslim). The PSI was implemented in half these schools between 2012 and 2016.

---

[4] The PSI is an initiative to decentralise school management by providing a mechanism such that key members from the local community (including parents) and schools can collaborate through the appointment of a school management committee. The PSI is expected to empower schools and build stronger partnerships with local communities facilitating fund raising opportunities, inter alia that can be used to improve student learning.

[5] Secondary schools can be type 'IAB', 'IC' or 2 in Sri Lanka. Type IAB offer curriculum streams science (including mathematics), arts and commerce that students can take for advanced level examinations at the end of grade 13. These schools teach grades 1-13. IC schools also teach grades 1-13 but offer only arts and commerce streams. Type 2 does not offer grade 12 or 13 classes. There are also small Type 3 schools that teach grades 1-5 or 1-8 but our sample does not contain any of these schools.

Around 20 students each were selected from the grade seven class of each of the 72 schools[7]. All schools are co-educational containing both boys and girls. Data collected for these children (the Student Survey) include test scores in Mathematics (with the same tests administered in both rounds of the survey) and their perceptions of the quality of their maths teacher as being accessible, friendly, encouraging class participation and giving regular feedback. The survey also included a 'Teacher Survey' collecting data from five teachers per school including the Mathematics and English teachers who taught the class of grade seven students in the Student Survey and three other teachers in language, Science, History/Geography and first language for the relevant Grade seven classes in each school. The questions asked were regarding their experience in teaching, motives for joining the profession, teaching styles and perceptions of class performance. In addition to this teacher survey a school-level survey was conducted with data gathered from the principal of the school regarding their own experience and qualification as well as school resources. Finally household level information for the children was gathered based on a separate household survey where the respondent was mainly the child's parent or parents. Each record for a student in the student survey corresponds to a unique household in the household survey (i.e., the sample contains one student per household). The same information was gathered from grade seven students in the same schools in 2016, their respective households, teachers and the principal after the PSI was implemented. Note that although the school

[6] There are five types of school 'levels': very congenial, congenial, uncongenial, difficult and very difficult as classified by the Ministry of Education in 2007 to depict the level of infrastructure and facilities of schools.

[7] If the class size exceeded 20, simple random sample method was used to select the students. However in a few schools there were less than 20 students in Grade 7.

remained the same giving the data its panel feature, the students, their households and teachers were different between the rounds.

Although not randomly assigned, Table 1 shows that PSI and non-PSI schools did not differ significantly in terms of baseline school characteristics. This is unsurprising given the efforts made to make sure that the 36 non-PSI schools were carefully selected to match the PSI-schools in terms of number of students, school type, level and race. The table also shows that there are no differences even in terms of most other individual and household characteristics. In the regression analysis in the next section, we include some of these variables directly into the regression so that they are no longer a part of the error term. This also reduces the chance that the coefficients on the treatment and time dummies are biased, in the event they are correlated with these variables.

[INSERT TABLE 1 HERE]

Table 2 provides summary statistics for the key variables used in the ensuing regressions, by gender, for the full sample of around 2600 children for whom complete data are available from both rounds of data. A significant difference lies in parental aspirations, favoring girls, regarding whether the child will pursue education beyond the compulsory schooling age of 14. This is likely to be influenced by a student's prior attainment as well as labor market opportunities and gender-stereotyping. Although the measure captures some of the implicit gender based biases and future opportunities it is crude not the least because 13.9 per cent of the parents (divided equally between boys and girls) were unsure about future plans and have been assigned category 1, arbitrarily, to avoid loss of observations. Excluding those who are 'unsure' from the sample leaves us with a sample of 1214 girls and 1072 boys, with mean aspirations 6.48 and 6.25 respectively, with a significant difference of 0.23 (standard error 0.09).

There is also a significant difference at the 10 per cent level with regard to how boys' rate their maths teachers compared to girls. The rating was based on answers to five questions in the student survey that asked the children how friendly and accessible teachers were within and outside class room, their perception on how knowledgeable the teacher is regarding the subject matter, the teacher giving regular feedback on work, using different types of methods to teach in class and encouraging class participation. A comparison of responses to these questions showed that girls score teachers significantly higher on accessibility, knowledge of subject matter and providing regular feedback than do boys. Although there is some evidence to suggest that boys' perception of teacher support may on average be lower than that of girls' (Wentzel, 1997), our data shows that teachers' own reports (based on the teacher survey) of the time they spend in typical 40 minute lesson explaining and providing feedback is negatively correlated with the proportion of boys in a class room. For example, in class rooms with more than 60 per cent boys, the average time spent providing feedback and using the blackboard to explain was 8 minutes while in a class with less than 40 per cent boys it was 10 minutes. The difference is significant at the 10 per cent level. Time spent explaining, getting students to read out the text book and class disciplining were all correlated negatively with the proportion of boys in a class, although not significant. Moreover, the time teachers spend on activities outside the class room unrelated to teaching during term time was significantly higher at 16 hours in a class with over 60 per cent boys while it was 9 hours when the proportion of boys was less than 40 (significant at the five per cent level). Thus there seems to be some evidence of biases in teacher behavior in class based on student gender that may be reflected in the student rating of teachers. So gender differences in this rating are considered to be indicative of implicit biases in the classroom favoring girls.

[INSERT TABLE 2 HERE]

The table also shows that significant differences between boys and girls exist in terms of time use: Boys tend to spend significantly more time playing with friends or watching TV compared to girls. This corresponds somewhat to teachers' observation in a case study conducted in Sri Lankan schools that suggests parental supervision is lower for boys (Aturupane, Shojo, & Ebenezer, 2018). Spending more time in 'leisure' may also be indicative of lower motivation and lower prior attainment. However, there are no differences in the hours spent on taking extra tuition (i.e., participating in shadow education). Children in the sample spent on average 3.5 hours taking extra tuition outside school hours in a typical week. The survey does not ask what subject extra tuition was taken in. Overall, child energy levels at school are likely to be low with 33 per cent of girls and 37 per cent of boys reporting not having had breakfast at least a few times over the past week. Although not reported in the table, around 20 per cent of the children also record that there were several days over the past year with no food at home for lunch after school. Finally, around 40 per cent of the maths teachers in the full sample are male. There is no significant difference in teacher experience.

Table 3 presents maths scores for boys and girls separately by year and treatment. For each year, the scores were standardized by subtracting the mean for non-PSI schools and then dividing it by the standard deviation for non-PSI schools. Thus the non-PSI group test scores have a mean zero and standard deviation of one for each year. Quite noticeably, these math scores are lower for boys compared to girls with differences significant, as seen in the last column[8]. The average scores for both boys and girls in all schools increased over time but the difference is only significant for girls in treated schools.

[INSERT TABLE 3 HERE]

3. Empirical Specification

To understand the effect of gender on maths scores in the 7th grade, we first estimate the following model that follows a standard difference-in-difference-in differences strategy (Angrist & Pischke, 2008; Lechner, 2011)

$$y_{ist} = \propto + \beta_1 X_{it} + \beta_2 year_{2016} + \beta_3 (PSI_s * year_{2016} * sex) + \beta_4 \boldsymbol{D}_s + \varepsilon \qquad (1)$$

where y_{is} refers to the standardised maths score of child i from school s during time period t (where t can either be 2012 or 2016). $\boldsymbol{X}$ is a vector of individual and household characteristics including gender (male=1), ethnicity, household income per capita (in 2012

[8] Another point to note is that scores are low overall for both boys and girls for a test marked on a 100 at 22.7 and 26.4 respectively, on average, pooling both rounds of data. This corresponds with several other studies that note how student attainment in Sri Lanka is low by international standards in spite of high enrolment (Dundar et al., 2017) .

prices) and the education level of the most educated in the household. It also includes gender interacted with $year_{2016}$. The time period dummy $year_{2016}$ equals 1 if the year is 2016 and 0 if it was 2012 captures changes in y over the two periods even in the absence of the intervention. PSI is a dummy that equals one if the school received treatment and zero otherwise and this is interacted with year and gender to capture the average treatment effect of the program on scores by gender. We also include school level fixed effects D_s, to account for characteristics that may be different between PSI and non-PSI schools. Standard errors are clustered by school to account for correlation of error terms within a given school. The coefficients of interest are the ones on the gender variable and the interaction variable. We also estimate (1) separately for boys and girls to see how the independent variables explain differently the correlates to maths scores between boys and girls[9].

To understand what drives gaps in attainment we use decomposition techniques. The first is the Oaxaca-Blinder technique that divides attainment between males and females into a part that is explained by differences in endowment (i.e., the determinants of attainment such as household income, parental aspirations, etc.), and a part that cannot be explained by such group differences (Jann, 2008). Thus the expected difference in attainment between boys and girls

$$A = E(y_{boys} - y_{girls}) \quad (2)$$

[9] The estimated equation for each group then becomes $y_{ist} = \propto + \beta_1 X_{it} + \beta_2 year_{2016} + \beta_3 (PSI_s * year_{2016}) + \beta_4 \boldsymbol{D}_s + \varepsilon$

where $E(y)$ denoting the expected value of the outcomes, can be accounted for by group differences in the predictors in (1), written below as

$$y_l = \gamma + \delta_l Z_l + \varepsilon_l, \qquad E(\varepsilon_l) = 0, \quad l \in \{boys,\ girls) \qquad (3)$$

where Z is a vector containing all the predictors in (1) apart from the gender dummy and a constant, δ contains the slope parameters and intercept. The mean outcome difference between boys b and girls g can be expressed as the difference in the linear prediction at the group specific means of the regressors:

$$A = E\left(y_b - y_g\right) = E(Z_b)\delta_b - E(Z_g)\delta_g \qquad (4)$$

Since

$$E(Y_l) = E(Z_l\delta_l + \varepsilon_l) = E(Z_l\delta_l) + E(\varepsilon_l) = E(Z_l)\delta_l$$

With $E(\delta_l) = \delta_l$ and $E(\varepsilon_l) = 0$ by assumption.

Re-arranging (4) as in Jann (2008)[10] gives a 'three-fold' decomposition of the outcome difference:

$$A = [E(Z_b) - E(Z_g)]\delta_g + E\left(Z_g\right)\left(\delta_b - \delta_g\right) + [E(Z_b) - E(Z_g)](\delta_b - \delta_g) \qquad (5)$$

[10] Jann (2008) cites Winsborough and Dickinson 1971, Jones and Kelly 1984 and Draymont and Andrisani (1984).

where the first summand on the right hand side of the equation refers the 'endowment' effect. This is the part of the differential that can be attributed to group differences in the explanatory variables. The second component refers to the 'coefficient' effect. This measures attainment differentials that can be attributed to differences in returns to the endowments. The last component is an interaction term accounting for the fact that differences in endowments and coefficients exist simultaneously. The fixed effects in the specification are explicitly accounted for as being fixed rather than stochastic to avoid inflation of standard errors. However, we also run the decomposition excluding the school fixed effects but including directly school specific characteristics such as type of school, principal characteristics to see whether it affects the results.

To check the robustness of results and generate confidence in our analysis we also apply unconditional quantile regression to decompose results at different points in the test score distribution (S. P. Firpo et al., 2018; Rios-Avila, 2020). The technique goes beyond simple mean comparisons, as in the Oaxaca-Blinder approach described above, to look at the entire density of scores. Thus the procedure allows us to examine whether the effect of the factors vary depending on the location in the distribution. This requires creating counterfactual densities rather than counterfactual means, had boys the same characteristics as girls. In the first stage, these counterfactual densities are created by using inverse probability weights and the decomposition conducted using the reweighted estimations (DiNardo, Fortin, & Lemieux, 1996). In our estimations the weights are calculated as follows. First the data for boys and girls is pooled to estimate a logit to find the probability of being male as a function of Z:

$$d_i = \theta + \vartheta Z_i + v_i \qquad (6)$$

where d is a dummy variable that equals 1 if child i is male and zero otherwise and vector Z comprises individual and household characteristics (income per capita, education level of the most educated, ethnicity, student rating of mathematics teacher), teacher characteristics (experience, gender) and school type. This is then used to predict the probability of being male $\widehat{Pr}(l_{male} = 1|Z)$, and being female $\widehat{Pr}(l_{male} = 0|Z) = 1 - \widehat{Pr}(l_{male} = 1|Z)$. The reweight factor $w(Z)$ for females is then calculated using the predicted probabilities and the sample proportion of males$\widehat{Pr}(l_{male} = 1)$ and females $\widehat{Pr}(l_{female} = 1)$, as follows:

$$w = \frac{\widehat{Pr}(l_{male} = 1|Z)/\widehat{Pr}(l_{male} = 1)}{\widehat{Pr}(l_{male} = 0|Z)/\widehat{\mathrm{Pr}}(l_{male} = 0)}$$

Identification relies on the key assumptions of ignorability and common support over the distribution of characteristics. The ignorability assumption states that the distribution of ε given Z in (3) is the same for boys and girls, though this does not mean that E(ε|Z)=0[11]. In the second stage, the endowment and coefficient effects are further disaggregated to reflect the contribution of each covariate using recentered influence functions (S. Firpo,

[11] The conditional independence assumption E(ε|Z)=0 can be replaced by the weaker ignorability assumption to compute the aggregate decomposition. For example, ε can be correlated with Z as long as the correlation is the same between males and females. Using language from the treatment effects literature the ignorability assumption can be restated as potential outcomes are independent of the treatment (i.e., being male or female), given a set of observables Z.

Fortin, & Lemieux, 2009) to estimate the effect that changes in Z have on test score percentiles 10, 30, 50 (median), 70 and 90[12].

4. Results

Table 4 columns 1-3 present results for (1) estimated together for boys and girls and for each gender separately using parsimonious 'specification 1'. Column 1 shows that gender matters significantly to scores with boys scoring less than girls, as expected, by about 0.24 on the standardized math score. Education level of most educated parent significantly influences math scores for both boys and girls but household per capita income affects positively only girls' scores. Columns 4-6 show results a less parsimonious with other variables included such as parental aspirations (entered as a categorical variable with aspire to complete up to compulsory education omitted), teacher attitude, teacher gender and teacher experience ('specification 2')- all identified in the literature as being potentially important. Before including these variables we checked first that these extra variables themselves are not affected by the treatment[13]. Including these extra variables mean that we lose 36 observations due to missing data on teachers or households. Nonetheless, the results confirm that boys score significantly lower than girls (column 4) and that the education level of the most educated in the household exerts a significant positive impact

[12] See Rios-Avila (2020) for further details on how unconditional quantile regressions can be estimated using recentered influence functions, and how this can be done within Stata, the software package employed in this paper.

[13] To check for this possibility we estimated specification 1 with parental aspiration, student rating of teacher, teacher experience and teacher gender in turn as dependent variables with PSI, PSI*year2016 and school fixed effects as independent variables. The unreported results confirm that the PSI had no impact on these variables

on the attainment of both genders. In addition, parent's aspiring that their child completes secondary school (grade 13) to pursue further vocational or university based education correlates with a significant positive effect on scores of both boys and girls. Teacher attitude is significant for girls' performance. Both specifications show that the PSI intervention has no impact on scores of either gender assuming that the parallel trend assumption (i.e., the difference between PSI and non-PSI schools is constant over time, in the absence of the intervention) holds. We test for this assumption and do not find evidence to suggest that the assumption has been violated (unreported). The result that the PSI has no impact on grade seven maths scores is similar to Aturupane et. al. (2014)'s finding in a previous run of the PSI during the 2006-2008 period in Sri Lanka that showed the PSI had no effect on any of the test scores of grade 8 students.

[INSERT TABLE 4 HERE]

The results do not support the notion that teacher gender is significant in explaining attainment. Counter intuitively, teacher experience (measured using years in service, ranging from 1 to 38 years) is significantly negatively correlated with boys' math scores. The coefficient on the square of teacher experience included to capture non-linearities- is insignificant. To verify robustness of this result we re-estimate the model replacing teacher experience and its square with teacher in service grade, a discrete variable ranging from one to five where grade one is associated with teachers who have a degree or have more training/experience and grade five is associated with teachers with new entrants and those with lower formal qualifications or training. Unreported results show that replacing teacher experience with categorical variable teacher grade renders the result that the 'better' is the teacher's grade, the poorer is the test score for both boys and girls. This maybe because in our sample of maths teachers, the less experienced (in-service categories four and five, 38 per cent of sample) seem to spend more hours a term on lesson planning, counselling, supporting extra-curricular activities and assessing work than the more experienced (in-service category one, 20 per cent of the sample), as seen in Figure 1. This is an aspect that needs to be investigated further.

The model estimated using other specifications (results unreported) including class room level variables pertaining to teaching style such as having clear incentive structures for the more able performers and safety nets for poorer performers (as reported by the teachers) showed that neither of these variables are significant[14].

[14] In yet some other specifications, we included variables such as time spent on leisure activities (e.g., watching TV) to note it has a significant negative impact on scores of both boys and girls. Replacing this variable with hours spent taking extra tuition shows that tuition exerts a positive significant effect on maths scores, again for both genders. As stated previously, the questionnaire does not ask the children as to the subject they receive extra tuition but it is likely that it is mostly in mathematics, science and English. It may also be possible that some of these variables such as hours spend in extra tuition in particular maybe affected by the treatment. We checked for this possibility by including all these additional variables in turn as outcome variables in specification 1 to find no significant average treatment effect. But including these additional variables may also be problematic if they are correlated with the error term. For example, unobservable such as ability, motivation and attitude may be correlated with both extra tuition and time spent in leisure and if that is the case the coefficients on these variables will be biased and we therefore do not report these results. Finally, we re-estimated tables 4 (columns 4, 5, 6), 5 and 6 using a sub-sample excluding observations where aspirations were recorded as being uncertain (14% of sample). As noted in Table 2, these aspirations for these observations were coded as 1 in our estimations. All results remain robust to the exclusion of these observations.

[INSERT FIGURE 1 HERE]

Decomposition Analysis

Table 5 presents the results of the Oaxaca-Blinder decomposition. The predicted female attainment is 0.20 while it is significantly lower for males -0.01. The decomposition results also show that if males had the same characteristics as females the consequent change in attainment is around 0.01, although statistically insignificant. This is unsurprising as the section 2 already showed how males and females do not differ significantly in terms of most observed contemporaneous household characteristics apart from parental aspirations and teacher rating. The detailed decomposition analysis (unreported) does confirm that aspirations are significant and that if boys had the same endowment of parental aspirations as girls' their scores will increase by 0.1 (significant at 5 per cent level). But these differences do not explain significantly the overall gap in attainment. The difference in attainment is explained almost entirely by the coefficient effect –or 'discrimination' as conventional decomposition literature would term it. Thus the change in male attainment if females' coefficients were applied to male characteristics is 0.25. The table also presents results for a two-fold decomposition that uses average coefficients over both groups (Reimers, 1983), assuming that 'discrimination'-positive or negative- is not specific to boys or girls. The results confirm those of the threefold analysis.

[INSERT TABLE 5 HERE]

To look at how the impact of the various characteristics across the distribution of mathematics scores, decomposition analysis for various percentiles are presented in Table

6. The results show that girls score higher than boys at all percentiles with the highest difference at the median where girls score 0.26 higher. Differences in characteristics between boys and girls do not seem to explain the total gap in scores at any of the percentiles apart from the highest end, the 90th percentile, when ethnicity being Sinhalese is important. In contrast, at all points in the distribution the differences are 'unexplained', matching the previous Oaxaca decomposition results. Further disaggregation of the unexplained component shows that if boys in the lower half of the score distribution faced the same coefficients to teacher attitudes as did girls, their scores would be significantly higher by 0.60, 0.93 and 0.70 at the 10th, 30th and 50th percentiles. Further disaggregation of the explained component shows that while differences in aspirations to complete a university degree and teacher attitudes are significant to scores of boys towards the middle of the distribution, the consequent increase in scores had they faced the same aspirations and teacher attitudes as girls is only about 0.01-0.03 standardized points. Thus differences in coefficients remain the main driver of differences in scores between boys and girls.

[INSERT TABLE 6 HERE]

In summary, the education level of the most educated in the household and parental aspirations are the most significant factors contributing positively to attainment of both boys and girls in our sample. However, for girls in particular, teacher attitudes are important. Although parental aspirations and teacher attitudes are lower for boys that they are for girls, these differences in endowments –using our imperfects proxies- don't explain total attainment differences. Almost all of the test score difference between boys and girls is unexplained. Put differently, it is due to differences in returns to endowments rather than the endowments per se.

So what may be driving the differences in returns to endowments? The literature on decomposition often attributes it to 'discrimination' and the next section looks beyond compulsory schooling to seek sources of possible discrimination that may have an indirect impact on attainment.

5. Looking beyond compulsory schooling

5.1. Enrollment beyond compulsory schooling and labor market asymmetries

This section looks at what maybe driving the lower enrollment of boys beyond the compulsory age of 14, compared to girls. It is likely that the key explanatory factors have a recursive effect on boys own attitudes to learning at younger ages. As can be seen in Figure 2, the proportion of women in full time education is higher amongst women than men at almost all ages between 15 and 25. Ages 15-17 corresponds to when a major public exam at the end of school grade 10 is undertaken (Ordinary Levels), age 18-19 corresponds to when Advanced Levels are taken (a major public exam at the end of school grade 12 that allows a student to qualify for university entrance), 20-23 correspond to the first few years at university or other higher education and 24-25 match, rather broadly, later years in tertiary education including the undergraduate honors level, masters or postgraduate studies. Enrollment beyond compulsory schooling can also include vocational education.

The data for this graph and the remainder of the analysis for this section comes from the Household Income Expenditure Survey carried out by the Department of Census and Statistics for the 2012/13 period. The sample is representative of the population, containing information gathered from 25,000 households from all the districts in the country including those in the North and East that had been excluded from the survey for nearly three decades due to the civil war from 1983-2009.

[INSERT FIGURE 2 HERE]

Table 7 shows that there are significant differences in enrollment disfavoring men particularly after age 20. The lower male enrollment particularly in age category 23-25 is not due to males having already completed the equivalent of a bachelor's degree, as the average years of completed education for this group is lower than that for females. On average, women in this age group have higher unemployment rates and lower labour force participation rates than men. However, labour force participation for those women who have completed secondary school or beyond is higher than for those without at 47 per cent versus 22 percent (unreported in the table). The comparable rates for men are 65 and 53 percent respectively. Also noticeable is that average monthly wage earnings from primary and secondary employment earning for positive earners (22 per cent of the sample) is significantly higher for men at Rupees 14 099.57 compared to that of women at Rupees 12,029. The pattern is the same with all earnings that include average monthly earnings through wage employment as well as net earnings from agricultural and non-agricultural activities (e.g., profits from self-employment): Men earn significantly higher than females. All monetary values are expressed in real terms adjusted for province-level spatial price variation. There are no statistically significant differences between men and women in terms of ethnicity, urban or rural location, household size and household wealth index[15].

---

[15] The wealth index is a composite measure of a household's cumulative living standards, and is thus not subject to issues of endogeneity as much as an indicator for the household's monthly income or consumption maybe. The wealth index is calculated as the simple average of three sub-indices: the housing quality index, the access to services index and the asset index. The housing quality sub-index reflects the welfare of household members in terms of housing-related comfort by looking at materials used in constructing the family's dwelling (walls, roof, and floor), and household density (number of rooms per person excluding the bathroom, kitchen and hallways). Equal weights are assigned assuming equal importance between indicators and the simple average of the four indicators in calculated. Wall quality is 1 if the walls are made of 'good' material such as brick, cement blocks or cabok (a laterite brick) and 0 otherwise; roof is 1 if made of asbestos, tiles or concrete and 0 otherwise; floor quality is 1 if material is cement, tiles or terrazzo and 0 otherwise. The rooms per person index is calculated as follows for household i, $\frac{(rooms_i - \min\_rooms)}{\max\_rooms - \min\_rooms}$ where min_rooms and max_rooms are the range of bed rooms in the dataset and $rooms_i$ refer to the number of rooms in the household. The access to services sub-index measures the household's ability to meet functional requirements of sound shelter. Again it is calculated as the simple average of four indicators: access to electricity (=1 if the household has electricity and 0 if not); access to safe drinking water (=1 if water comes from protected wells or piped from a mainline and 0 otherwise); access to a safely managed sanitation.

Service (=1 if the type of toilet used is connected with water seal and 0 otherwise) and access to adequate fuel for cooking (=1 if gas, kerosene or electricity). All four indicators are considered to have equal weight. The consumer durables sub-index is a measure of the household's ownership of common household durable goods and assets (such as TV, radio, bicycle, cars). The index counts a long list of items (regardless of its monetary value) that the household has ($assets_i$) and converts it into and index as follows: $\frac{\sum assets_i - assets\_min}{assets\_max - assets\_min}$ where $assets\_min$ and $assassets\_max$ is the range of these items in the dataset.

The table also reports average predicted wages for 15-25 year olds. These predicted values are used as a proxy for opportunity cost of schooling in the regression estimation below. This measure for opportunity cost attempts capture what an individual will earn if he or she were not enrolled at school[16]. To calculate predicted wages for each individual male, we first estimated a multivariate regression equation of wages on individual characteristics that included age group (15-17; 18-22 or 23-25), current education level (completed compulsory years; completed Ordinary Levels; completed Advanced Levels or has post-school qualification), location (rural or urban) and province of residence for those men in the labour force (i.e., employed and unemployed, 54 per cent of the sample). Note that those employed recorded a positive wage while those unemployed recorded zero wages. The estimated equation is used to calculate predicted wages for those men in the labour force. For those not in the labour force (i.e., students), the predicted wage is the average predicted wage for the subgroup the student belongs to by province, location, education level and age group. For females, due to potential sample selection issues with regard to labour force participation, predicted wages were estimated using the Heckman correction for selection (Heckman, 1976, 1979). Thus we used data for employed women who recorded a positive wage (17 per cent of the sample), unemployed who recorded zero wages (10 per cent of the sample) and those whose labour force status was 'household work' (27 per cent of the sample) whose wage values are 'missing'. We then use the Heckman two-step procedure. In the first step the selection equation includes marital status as well as age level, education level, location and province of residence. The second step outcome model includes the same variables apart from marital status. The output was then

[16] Authors thank an anonymous referee for this suggestion while acknowledging as the referee did, that it is not perfect.

used to predict wages for those employed, unemployed and in household work. Predicted wage for female students was estimated as the average of the predicted wage for women in the relevant sub-group by province, location, education and age group. Those individuals whose labour force status was 'unable to work' (0.7 per cent of all 15-25 year olds) were assigned a predicted wage of 0. Opportunity cost is difficult to measure and we acknowledge that our estimations are imperfect for several reasons. For example the costs of obtaining education are excluded. Although education from primary right up to (and including university) is 'free' in Sri Lanka for those pursuing state-provided education, there are some private institutions providing in particular tertiary level education. There are also other costs in terms of private tuition, travelling and equipment that will have to be borne. Moreover, our estimations assume that individuals form expectations about earnings forgone adaptively, with current wages influencing their decisions. To make sure that our results are robust, we include an alternative measure of opportunity cost replacing wages with earnings to calculate predicted earnings, using the same methodology as described above.

[INSERT TABLE 7 HERE]

The following model is estimated to understand more about what drives an individual to invest an extra year in formal education beyond compulsory education:

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon \qquad (7)$$

where y refers to the enrollment status of individual i in post compulsory education between ages 15 and 25 (=1 if the individual is enrolled and 0 otherwise), X_1 is a vector of individual and household characteristics such as gender, age (in categories 18-19, 20-23, 24-25 with 15-17 omitted), ethnicity (Sinhala, Tamil with Muslim and other minorities omitted), marital status (=1 if never married and 0 otherwise), household size, urban residence and household wealth index that may all influence the schooling decision and X_2 is a vector of variables measuring school quality (proxied imperfectly by district-level indicators of school quality such as the experience of teachers and the number of IAC schools[17]). X_3 is a vector containing the opportunity cost of taking up an year in education proxied by the earnings forgone, as explained in the previous paragraph. It also contains categorical dummies for the province of residence. To see if factors such as household wealth, opportunity costs and school quality affect males differently to how it does females, the specification also includes several interaction terms, with each of these variables interacted with 'male'.

The probit estimation output for (7) is presented in Appendix Table A1 and the corresponding average marginal effects for selected variables are presented in Table 8a. Results in Column (1) are based on a specification that uses predicted wages as a proxy for opportunity cost while the specification in column (2) uses predicted earnings. The results are very similar under both specifications and for ease of exposition we focus on (1). As the first column shows, on average the probability of a male being enrolled is 0.05 smaller than that of a female. In terms of other covariates, household wealth index increasing by a unit

[17] This data is taken from the Ministry of Education's data on 'Teachers in national schools by age, 2013' http://www.moe.gov.lk/english/images/Statistics/sri_lanka_education_information_2013.pdf. Teacher experience is proxied by the proportion of teachers over 41 years of age, assuming older teachers have more experience (note that a majority of teachers in national schools enter the profession soon after leaving school).

increases the probability of being enrolled by 0.56, opportunity cost increasing by 1000 Rupees decreases enrollment by 0.01, household size increasing by a unit reduces enrollment by 0.01. The individual being Sinhalese in ethnicity rather than a minority group increases the probability by 0.07. Being never married rather than married, divorced or separated increases enrollment probability by 0.31. Age has been entered in the specification as a categorical variable and shows that there are significant falls in the probability of being enrolled as age category increases compared to ages 15-16.

[INSERT TABLE 8 HERE]

Do any of these factors matter significantly more for males than it does for females? Further analysis of the results looking at "interaction effects" observed through pairwise comparison of average marginal effects reported in Table 8b show that the average marginal effect of wealth on enrolment is 0.12 higher for males compared to females suggesting that the poorer is a household the more likely it is for males to drop out of further education. The average marginal effect of school type being 1AB on enrollment is 0.53 higher for males compared to females. This suggests a strong subject preference biased towards mathematics and sciences in male enrolment decisions beyond compulsory schooling. We also find that foregone earnings as a result of engaging an extra year at school matters significantly more to males than it does to females with a higher average marginal effect of opportunity cost of 0.01.

All the average effects discussed above are marginal (partial) effects calculated for each observation in the data and then averaged. As an alternative we obtain effects for male at the mean values of the variables with which 'male' is interacted (i.e., wealth, opportunity cost and the school variables). The effects are identical to those of the previous discussion. We also obtain effects for male at mean values for all independent variables. The results show that the probability of a male being enrolled is 0.08 smaller than that of a female (standard error 0.011). Pairwise comparisons of average marginal effects show that the effect of wealth is 0.19 higher for males than for females and the effect of school type being 1AB is 0.9 higher for males compared to females. Moreover, the marginal effect of opportunity cost at means suggests that increasing by 1000 Rupees decreases enrollment by 2 per cent. The results are robust to the exclusion of the school-level variables (unreported).

We also re-estimated (7) for the subsample of those who have completed secondary education (i.e., passed Advanced Level or equivalent) and are college-going age (20-25). This is 13.2 per cent of the full sample. The probit results are presented in Appendix Table A1 columns 3 and 4 that use predicted wages and predicted earnings, respectively, to estimate opportunity cost. The corresponding average marginal effects in Table 8a columns 3 and 4 show that the opportunity cost increasing by 1000 Rupees decreases enrollment by 2 per cent. However, gender is not significant. Moreover the pairwise comparison of average marginal effects (unreported) do not indicate that opportunity cost is significantly more important to males compared to females. Thus once youth have invested 5 extra years beyond compulsory schooling and passed the grade 13 Advanced level exams that qualify them for university-level qualifications, gender effects stop being significant, *ceteris paribus*.

A key conclusion from the above analysis, important to our narrative, is that for men, particularly those who have not completed secondary education who comprise over 80 per cent of our sample, the opportunity cost of education has a higher significant negative bearing on the decision to pursue an extra year of education compared to women. This may be due to societal norms (i.e., a higher expectation for men to contribute to household incomes rather than women). But it may also be because for every occupational category, the average earning for men is significantly higher than those for women, apart from the managerial category (Table 9, columns 1-3). Earnings include employment earnings as well as net earnings from agricultural and non-agricultural activity (excluding in kind earnings, transfers and remittances) for those between 15 and 30 years of age. We replaced earnings with total incomes to find that the significance of gaps between men and women remain the same for all categories apart from 'sales'.

[INSERT TABLE 9 HERE]

Table 9, columns (4)-(6) also show that for every occupational category, the average education level of men is significantly lower than those of women. This maybe reflective of men having a lower 'taste' for education or positive discrimination in the labor market where men have to be less educated than women, on average, to secure the same job. Himaz (2010) argues that it is incorrect to suppose that women have a higher 'taste' for education as pro-female biases in within household education expenditure occur only close to key public exams. If acquiring higher level of education was down to taste, then more should be spent on girls at every stage of education. Thus girls acquiring more education –or boys acquiring less- is unlikely to be down to 'taste'. It is more likely that there is discrimination in the labor market that requires women to acquire more qualifications in order to vie for the same role as a man. There is also evidence that returns to education are quite low for women who only complete compulsory education compared to men, while it is much higher for higher levels of education compared to the returns for men for comparative grade[18]. This again maybe another reason as to why parental aspirations are lower for men than women.

Another trend observable in Table 9 columns (7)-(9) is that over half of young employed women aged 15-30 work in occupations strongly associated with formal education such as professional and technical occupations (e.g., teachers, doctors, engineers, and draughts persons), clerical support and sales related jobs. Only a third of young men

---

[18] For example, using the same dataset we estimate labor market returns for men and women separately. We find that for men, education categories compulsory, passed O-levels, passed A-levels and degree have a return of 1.1, 4.1, 7.8 and 5.7 per cent while for women it is 0.8, 9.1, 14.4 and 10.6 per cent, compared to having no education. To estimate returns we applied fixed effects estimation to a cluster sample (Behrman & Deolalikar, 1993, 1995).The cluster in this case is the household and the sample is limited to households that have two or more males (females) who are in wage employment.: The method accounts for sample selection indirectly through correcting for unobserved heterogeneity at the household level, in so far as the heterogeneity bias results from attributes that are common to individuals in the same household. Thus we estimate the following separately for males and females between ages 18-55: $y_{ij} = \beta_0 + \beta_1 X_{ij} + \sum_{k=0}^{4} \beta_2 s_{ijk} + \beta_3 f_j + \varepsilon$ where y is log of earnings in real terms from primary and secondary employment for individual i in household j, X is a vector of individual and household characteristics (age, square of age, wealth index, household size, rural residence, ethnicity and province of residence) and $s_0 \dots s_4$ are dummy variables reflecting level of education numbered in the following order:0=no schooling, 1=compulsory education(9 years), 2= passed O-levels (11 years), 3=passed-A-levels (13 years) and 4=gained a post-school qualification (a degree, professional or vocational qualification after reading for a-levels, roughly 16 years), with no schooling omitted in the estimations. Unobserved fixed effects are captured by f. The return to education level k is estimated as $(\beta_k - \beta_{k-1})/(s_k - s_{k-1})$. The number of men used in this restricted sample is 2165 coming from 1007 households and 835 women coming from 392 households. Replacing y_{ij} with employment earnings plus net earnings from agricultural and non-agricultural activities also produces returns similar in patterns.

are engaged in this type of occupation. The professional roles arguably offer better working conditions, job security (particularly if employed in the state sector) and benefits in kind (such as more holidays if employed in teaching although the remuneration is low). Nearly 70 per cent of young men are engaged in skilled agriculture and fisheries, plant and machine operator roles (a third of who are bus or three-wheeler drivers, roughly 5 per cent of all employed males aged 15-30), craft and related roles (building and related works, electrical and electronic trades, etc.) and elementary occupations. Thus significantly fewer men are employed in occupations in the managerial, professional, technical, clerical and sales categories. This may be due to men having acquired lower education than women and therefore being crowded out of the market for professionals and technicians or that men have a preference towards occupations that require lower schooling. Or it may be that for men, average incomes are higher than those for women- so much so that completing even less than compulsory education (8.4 years of schooling) can bring an average earning of Rs. 12850.7 working in elementary occupations while for women to earn on a similar amount of Rs. 12347.9 on average in a sales related occupation requires an education level of 10.2 years (i.e., has completed ordinary levels). This observation is supported by (Gunawardena, 2015) who notes that males are treated more favorably in the labor market in spite of similar non-cognitive skills but lower measured cognitive skills compared to women, particularly in the 20-29 group . Not only are men less unemployed, they also engage in a wider set of occupations- indicative of a wider choice- compared to women (Appendix Figure A4).

Thus wide differentials in how the labor market values education between men and women may be at the root to why attitudes to learning and returns to endowments are significantly different between boys and girls. Possible contributors to this are supply-side bottle necks in higher education. Places in state universities, for instance, are highly competitive and only 17 per cent of those who qualified for entrance via the Advanced Level examination went on to being admitted to university in 2013 . In spite of this tough competition to secure a place at university, the monetary return to tertiary education is low with 70 per cent of university graduates employed in the government or semi-government sector, over 60 per cent of whom take up teaching. The jobs offer security and other benefits but are low in terms of monetary returns. Those courses in higher education that enjoy a higher return such as IT or medical degrees face severe supply-side bottlenecks and when offered by the private sector can be prohibitively expensive. It is likely that these bottlenecks and constraints in higher education, relative return structures as well as labor market asymmetries all play a strong role in causing returns to endowments to vary significantly by gender at the school level, and be realized in terms of lower attainment for boys compared to girls.

5.2. Education as a substitute for the dowry at marriage

Another possible explanation for the female bias in education attainment and parental aspirations is the role the marriage market may play. For example, if education were a substitute for the dowry at marriage -- an endowment by parents to their daughter and son-in-law at the time of marriage in the form of money or assets - parents may choose to invest in girls' education rather than the dowry. Indeed, the more educated is a woman, the higher is her labour force participation rate, in the case of Sri Lanka. For example, if we consider the age group 30-49 that comprises mostly those who are beyond the usual age of being engaged in formal education, the differences in labour force participation by education level are stark. As seen in Table 10, around 92.7 percent of women educated beyond secondary school level (i.e., have a university degree, diploma or professional qualification) participate in the labour force compared to 48.7 per cent who have completed secondary school (i.e., passed Advanced Level or equivalent) and 37.8 of those who did not complete secondary school. For men, however, labour force participation remains high at over 97 per cent irrespective of the education level attained. Thus women's labour force participation is strongly correlated to education. Given that the median age of first marriage in Sri Lanka among 30-49 year old ever married women in 2016 was 28.7 years for those with a degree and 26.3 for those who completed secondary school compared to 20-23 years for those who have not completed secondary education(DHS, 2016, p. 83), it appears that females have at least a few years to engage in the labour market to collect their own dowries before marriage. Can this explain, at least in part, as to why parents aspire that daughters attain higher levels of education and girls themselves are keen on performing better at school compared to boys? Evidence does not support this contention. If more education and later marriage meant girls collect a bigger dowry to facilitate marriage then higher percentages of better educated women should be married, compared to women with lower education. But as the table shows, 11.9 per cent of women aged 30-49 educated beyond secondary school have never married (legally or according to custom/repute) compared to 5.2 who are educated to a level less than secondary school. So the connection between education, dowry accumulation through employment and marriage are not clear, matching observations in Andersen (2007). What the lower rates of marriage among more educated women does indicate, however, is that the lower acquiring of education on the part of men may have implications for relationships and family formation as women are less likely than men to form families with a less educated spouse (Grow & Van Bavel, 2015; Schwartz & Han, 2014; Therborn, 2004). Corroborating the latter point is the higher tendency for positive assortative mating by level of education seen among married couples in the upper wealth quintiles in Sri Lanka (Himaz & Aturupana, 2018, pp. 595-596).

[INSERT TABLE 10 HERE]

It has also been noted that the importance of dowries at marriage is shrinking with only around 20 per cent of youth considering it necessary in a 2002 survey of Sri Lankan youth (Hettige & Mayer, 2002). There are, however, class-based differences with regard to dowries with those from higher socioeconomic classes being able to afford to give dowries while low and lower middle income household struggling to do so, relying on education and employment[19]. Thus for poorer households in particular, potential future earnings that support a household are likely to be a higher concern than dowry, when deciding to invest in girls education, particularly in the context of slow economic growth and political unrest experienced in the past few decades. Moreover, a daughter's earnings after marriage do not accrue mainly to the family of her in-laws in Sri Lanka with her natal family being able to benefit as well. Unlike in neighbouring India and Bangladesh, Sri Lanka's cultural heritage indicates relative gender equality in terms of bilateral descent, a daughter's value within the natal home and continued support from natal family even after marriage (Malhotra & Tsui, 1996, p. 479). Thus economic returns to education, opportunity costs and indeed the probability of labour force participation seem to be what drive household decisions to invest in a child's education more than dowry or marriage market related concerns.

[19] Many thanks to Prof. Siri Hettige for raising this point via personal correspondence.

6. Conclusion

This paper looked at whether the lower attainment in math scores among boys in grade 7 in Sri Lanka can be explained by boys responding significantly differently than do girls to parental aspirations, teacher attitudes, teacher gender, household characteristics and decentralized, school-based management. The paper found that the attainment of both boys and girls were affected significantly by the education level of the most educated, parental aspirations and teacher experience (but not gender), with girls particularly sensitive to teacher attitude (i.e., a child's rating of how accessible and supportive the teacher is). On average, parental aspirations and teacher attitudes, albeit imperfectly proxied, were significantly lower for boys than for girls. Thus there are implicit biases disfavoring boys. But decomposition analyses showed that differences in endowments explained almost none of the difference in scores between boys and girls. Instead, the differences were 'unexplained', driven by returns to these endowments. In the decomposition literature, this is often attributed to 'discrimination'. In order to investigate the source for this 'discrimination', against boys as it would seem, we looked for explanations beyond compulsory schooling.

The second result of the paper noted that males were more sensitive to foregone earnings in their decision to participate in education beyond compulsory years, even after controlling for household wealth. Men earn on average more than women in Sri Lanka for every level of education in every occupational category and have a wider choice of occupations to engage in compared to women. Moreover, state university places are competitive, limited and often offer low monetary returns, while private sector higher education options are expensive. So the time, effort and resources required to secure a place in higher education maybe seen as less preferable to seeking employment soon after compulsory education especially for men who have higher forgone earnings. Thus asymmetries in the labor market favoring men and bottle necks in higher education may all have a role in creating a situation where the returns to endowments are lower for men at the school level that is realized in lower attainment. Put differently, boys, their parents and teachers realize that boys can put in less effort and attain less at school but still enjoy labor market outcomes that are more favorable than those experienced by their female counterparts. The source for the unexplained element of the decomposition result for school-level attainment disfavoring boys, therefore, may lie partly in positive discrimination favoring boys in the labor market. Although these aspects need further research a strong implication then is that the differences in attainment and enrollment may well be a rational response to asymmetries in the labor market and higher education system.

So if men are positively discriminated in the labor market, does the poorer attainment and lower enrollment levels of men matter in the long run? We believe is matters very much as the disparity is a symptom of deeper structural problem and is as important as it would be if the trends were reversed. At one level, having to acquire more years of education than men to end up earning less is in itself discriminatory for women. But at a broader level, a breakdown of male occupations showed that in spite of being employed, such employment is mainly in low or semi-skilled work. The lower accumulation of numerical and other skills on the part of males inhibits their hierarchical movement in the occupational ladder and taking up more productive roles. The disparity is also indicative of a mismatch in the labor market between what is demanded by employers and what is supplied by the education system that males respond to by acquiring less human capital than what maybe socially optimal. It may also reflect the fact that talent allocation in the labor market is not based on education either due to corruption or due to education taking on more a signaling role rather than a productivity enhancing role, and gender based reactions to this are different due to asymmetries in the labor market. Thus a key insight our analysis provides is the importance of looking beyond simply school-based reforms to improve more gender parity in attainment and enrollment. The Sri Lankan case suggests the importance of higher education and labor market policy reforms are needed to address the issue of gender imbalances in attainment.

Quite apart from labor market related aspects, the widening gap is also a grave concern for the destabilizing effects the widening horizontal inequities can entail such as unrest, war, frustration, corruption. There could also be important implications for relationships and family formation as women are less likely than men to form families with a less educated spouse. Overall there may be important risks associated with males' disadvantage in educational attainment just as there were disadvantages associated with females' lower participation in the past. These aspects also warrant further investigation.

References

Abdullah, M. C. (2009). Making gender sense in Malaysian secondary schools. In E. Page & J. Jha (Eds.), *Exploring the bias: Gender and stereotyping in secondary schools* (pp. 105-140). London: Commonwealth Secretariat.

Anderson, S. (2007). The Economics of Dowry and Brideprice. *Journal of Economic Perspectives, 21*, 23.

Angrist, J. D., & Pischke, J. S. (2008). M*ostly harmless econometrics: An empiricist's companion*. In: Princeton University Press.

Aturupane, H., Glewwe, P., Ravina, R., Sonnadara, U., & Wisniewski, S. (2014). An Assessment of the Impacts of Sri Lanka's Programme for School Improvement and School Report Card Programme on Students' Academic Progress. *Journal of Development Studies, 50*(12), 1647-1669. doi:10.1080/00220388.2014.936396

Aturupane, H., Shojo, M., & Ebenezer, R. (2018). Gender Dimensions of Education Access and Achievement in Sri Lanka. In: World Bank.

Behrman, J. R., & Deolalikar, A. B. (1993). UNOBSERVED HOUSEHOLD AND COMMUNITY HETEROGENEITY AND THE LABOR-MARKET IMPACT OF SCHOOLING - A CASE-STUDY FOR INDONESIA. *Economic Development and Cultural Change, 41*(3), 461-488. doi:10.1086/452028

Behrman, J. R., & Deolalikar, A. B. (1995). ARE THERE DIFFERENTIAL RETURNS TO SCHOOLING BY GENDER - THE CASE OF INDONESIAN LABOR-MARKETS. *Oxford Bulletin of Economics and Statistics, 57*(1), 97-117. doi:10.1111/j.1468-0084.1995.tb00029.x

Cornwell, C., Mustard, D. B., & Van Parys, J. (2013). Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments Evidence from Primary School. *Journal of Human Resources, 48*(1), 236-264. doi:10.1353/jhr.2013.0002

DHS. (2016). Demographic and Health Survey, Sri Lanka, 2016: Final Report. In. Department of Census and Statistics, Sri Lanka.

DiNardo, J., Fortin, N. M., & Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica, 64*(5), 1001-1044. doi:10.2307/2171954

Dundar, H., Millot, B., Riboud, M., Shojo, M., Aturupane, H., Goyal, S., & Raju, D. (2017). *Sri Lanka education sector assessment :Achievements, challenges, and policy options (English)* . . In *Directions in development; human development.* . Washington. D. C.: World Bank Group.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica, 77*(3), 953-973. doi:10.3982/ecta6822

Firpo, S. P., Fortin, N. M., & Lemieux, T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics, 6*(2). doi:10.3390/econometrics6020028

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions. R*esearch Study Brief.* . In. New Haven, CT: Yale University, Yale Child Study Center.

Grow, A., & Van Bavel, J. (2015). Assortative Mating and the Reversal of Gender Inequality in Education in Europe: An Agent-Based Model. *Plos One, 10*(6). doi:10.1371/journal.pone.0127806
Gunawardena, D. (2015). *Why Aren't Sri Lankan Women Translating Their Educational Gains Into Workforce Advantages?* . In. Washington, DC.: Center for Universal Education at Brookings.
Heckman, J. J. (1976). COMMON STRUCTURE OF STATISTICAL-MODELS OF TRUNCATION, SAMPLE SELECTION AND LIMITED DEPENDENT VARIABLES AND A SIMPLE ESTIMATOR FOR SUCH MODELS. *Annals of Economic and Social Measurement, 5*(4), 475-492.
Heckman, J. J. (1979). SAMPLE SELECTION BIAS AS A SPECIFICATION ERROR. *Econometrica, 47*(1), 153-161. doi:10.2307/1912352
Hettige, S., & Mayer, M. (2002). *Sri Lanka Youth, Challenges and Responses* . In. Colombo.: Friedrich Ebert Stiftung.
Hillman, N., & Robinson, N. (2016). *Boys to Men: The underachievement of young men in higher education-and how to start tackling it*. In. Oxford: Higher Education Policy Institute.
Himaz, R. (2010). Intrahousehold Allocation of Education Expenditure: The Case of Sri Lanka. *Economic Development and Cultural Change, 58*(2), 231-258. doi:10.1086/648187
Himaz, R., & Aturupana, H. (2018). Schooling and household welfare: The case of Sri Lanka from 1990 to 2006. *Review of Development Economics, 22*(2), 592-609. doi:10.1111/rode.12355
Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *Stata Journal, 8*(4), 453-479. doi:10.1177/1536867x0800800401
Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics, 4*(3), 59.
Malhotra, A., & Tsui, A. O. (1996). Marriage timing in Sri Lanka: The role of modern norms and ideas. . *Journal of Marriage and the Family , pp.476-49, 58*, 14.
Reimers, C. W. (1983). LABOR-MARKET DISCRIMINATION AGAINST HISPANIC AND BLACK-MEN. *Review of Economics and Statistics, 65*(4), 570-579. doi:10.2307/1935925
Rios-Avila, F. (2020). Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition. *Stata Journal, 20*(1), 51-94. doi:10.1177/1536867x20909690
Schore, A. N. (2017). ALL OUR SONS: THE DEVELOPMENTAL NEUROBIOLOGY AND NEUROENDOCRINOLOGY OF BOYS AT RISK. *Infant Mental Health Journal, 38*(1), 15-52. doi:10.1002/imhj.21616
Schwartz, C. R., & Han, H. Y. (2014). The Reversal of the Gender Gap in Education and Trends in Marital Dissolution. *American Sociological Review, 79*(4), 605-629. doi:10.1177/0003122414539682
Therborn, G. (2004). *Between sex and power: Family in the world 1900-2000*. In: Routledge.
UNESCO. (2015). Global Monitoring Report: Education For All 2000-2015: Achievements and Challenges. In: United Nations Educational, Scientific and Cultural Organisation.
UNGEI. (2012). *Annual Report*. Retrieved from http://www.ungei.org/FINAL_Annual_Report_2012.pdf
UNICEF. (2006). *Why are Boys Underperforming in Education? Gender Analysis of Four*

Asia-Pacific Countries. Retrieved from http://www.ungei.org/report_why_are_boys_underperforming_FINAL.pdf

Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology, 89*(3), 411-419. doi:10.1037/0022-0663.89.3.411

Wongmonta, S., & Glewwe, P. (2017). An analysis of gender differences in household education expenditure: the case of Thailand. *Education Economics, 25*(2), 183-204. doi:10.1080/09645292.2016.1168363

Younger, M., Warrington, M., & McLellan, R. (2005). *Raising boys' achievement in secondary schools: Issues, dilemmas and opportunities* In. UK: McGraw-Hill Education

Table 1: Selected baseline characteristics between PSI and non-PSI schools

	Non-PSI (1)	PSI (2)	Difference in means (1)-(2)
School Characteristics			
Offers Science (Type 1AB)	0.17	0.19	-0.03 (0.09)
Offers only arts/commerce (Type C)	0.50	0.58	-0.08 (0.12)
Classes only up to grade 11 (Type 2)	0.31	0.22	0.10 (0.10)
Principal's professional grade[1]	3.45	3.27	0.18(0.19)
Male Principal (proportion)	0.82	0.88	-0.06(0.08)
Proportion of males in grade 7	0.45	0.46	0.00 (0.04)
Observations (schools)	36	36	
Student Characteristics			
Missed days of school	13.08	12.08	0.80(0.57)
Time play (minutes per week)	303.03	291.63	11.39(6.30)*
No breakfast (proportion)	0.43	0.25	0.17***(0.02)
Co-curricular activities (proportion)	0.47	0.44	0.06***(0.01)
Household Characteristics			
Most educated (years)	9.56	9.38	0.17(0.15)
Household size (log))	1.59	1.59	-0.00(0.01)
Income (log per capita)	6.32	6.82	0.04(0.06)
Observations (students/households)	664	682	

Notes: [1]Ranges from 1 (certified teacher) to 5(Ph.D. or MPhil in education). Standard errors in parenthesis. Significance of standard errors *** p<0.01, ** p<0.05, * p<0.1

Table 2: Mean Individual and household characteristics by gender for pooled sample of grade 7 students (2012 and 2016)

Variable name	Variable description	Girls	Boys	Differences in means (1)-(2)
Individual and Household characteristics				
Parental aspiration	Response to question asked from parent: 'What is your child's plan after Compulsory (Basic) Education?': 1=stop after compulsory education, 2=compulsory education + vocational course, 3=Complete grade 11, 4=complete grade 11+vocational course, 5= complete grade 13, 6=complete grade 13+vocational, 7=complete a professional course, 8=complete a degree. Those respondents who say they are 'unsure' (13.9%) want to start their own business (0.9%) or do not intend working (0.3%) have been coded as 1[1].	5.77	5.46	0.31**(0.11)
Teacher rating	Average of the student's rating of his or her maths teacher ranging from 1 (strongly disagree) to 5 (strongly agree) on each of the following five questions: Teacher is very friendly and approachable in class and outside; Teacher uses different methods to teach (e.g. Discussions, group work, field visits); Teacher encourages group activities in class (discussion/projects); Teacher seems to have a good knowledge about the subject;	4.09	4.06	0.03* (0.02)

	Teacher gives you feedback on your assessments.			
Income	Log of total monthly household income per household member (2012 prices)	6.35	6.50	-0.15***(0.04)
mosteduc	Education level of the most educated parent in the family, ranging from 0 years (no education) to 16 years (Ph.D. or post graduate study)	9.44	9.68	-0.24*(0.11)
Sinhala	=1 if ethnicity of household head is Sinhala, 0 otherwise	0.64	0.67	-0.03(0.02)
Tamil	=1 if ethnicity of household head is Tamil, 0 otherwise	0.20	0.23	-0.03(0.02)
Time play	Total time (in minutes) a student spends in a typical week on playing on his/her own or with friends and on TV/computer entertainment.	269.5	302.6	-33.09***(4.45)
Tuition hours	Average number of hours per week during the school year spent attending tuition classes outside school hours	3.42	3.41	0.01(0.13)
No breakfast	=1 if student’s response is ‘yes’ to ‘During past week were there several days you did not have a breakfast?’, 0 otherwise	0.33	0.38	-0.0516**(0.02)
Mathematics teacher characteristics				
Teacher gender	Gender of maths teacher (=1 if male)	0.42	0.46	-0.03(0.02)
Teacher	(1) In service grade of teacher (1=highest grade, 5=lowest grade)	2.92	2.88	0.04(0.05)

experience				
	(2) Years in service	12.82	12.61	0.21(0.37)
Number of Obs.	Total number of grade 7 students from the 2012 and 2016 surveys with complete information on test scores, teacher and household characteristics.	1395	1261	

Standard errors in parenthesis. Significance of standard errors *** p<0.01, ** p<0.05, * p<0.1

Table 3: Standardised math scores by age and year

	Boys		Girls		
	N	Mean (1)	N	Mean (2)	Difference in mean (1)-(2)
PSI Schools					
2012	316	-0.01	366	0.21	-0.20***(0.08)
2016	323	0.12	321	0.37	-0.24***(0.08)
Change over time		0.11(0.08)		0.16**(0.08)	
Non-PSI Schools					
2012	305	-0.11	359	0.10	-0.21***(0.08)
2016	317	-0.08	349	0.12	-0.21***(0.08)
Change over time		-0.03(0.08)		0.03(0.08)	
All Schools					
2012	621	-0.05	725	0.15	-0.20***(0.06)
2016	640	0.02	670	0.24	-0.22***(0.06)
Change over time		0.07(0.06)		0.09(0.06)	

Significance of standard errors (in parenthesis), *** p<0.01, ** p<0.05, * p<0.1

Table 4: Grade 7 math scores (school level fixed effects with clustered standard errors). *Dependent variable: Standardised maths score*

	(1)	(2)	(3)	(4)	(5)	(6)
male	-0.24***			-0.21**		
	(0.09)			(0.09)		
y2016*psi*male	-0.03			0.02		
	(0.18)			(0.18)		
psi*male	-0.07			-0.11		
	(0.11)			(0.11)		
y2016*psi	0.09	0.03	0.06	-0.01	0.01	-0.06
	(0.17)	(0.15)	(0.18)	(0.17)	(0.15)	(0.18)
income	0.02	-0.01	0.05*	0.02	-0.01	0.03
	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)
mosteduc	0.08***	0.07***	0.08***	0.07***	0.06***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Sinhala	-0.11	-0.05	-0.07	-0.14	-0.04	-0.16
	(0.14)	(0.21)	(0.22)	(0.14)	(0.21)	(0.20)
Tamil	-0.23*	-0.20	-0.19	-0.28**	-0.20	-0.28**
	(0.13)	(0.23)	(0.14)	(0.11)	(0.22)	(0.13)
Parental Aspirations						
Compuls+vocational				0.18*	0.17	0.20
				(0.10)	(0.17)	(0.13)
Complete grade 11				-0.05	-0.11	0.06
				(0.11)	(0.14)	(0.14)
Grade 11+vocational				-0.03	-0.24*	0.11

				(0.12)	(0.14)	(0.16)
Complete grade 13				0.13	0.13	0.15
				(0.08)	(0.10)	(0.12)
Grade 13+vocational				0.18**	0.17*	0.24**
				(0.08)	(0.10)	(0.10)
Grade 13+Professionl				0.15**	0.14	0.16
				(0.07)	(0.11)	(0.11)
University degree				0.29***	0.26***	0.33***
				(0.06)	(0.08)	(0.07)
Teacher rating				0.09**	0.04	0.15***
				(0.04)	(0.05)	(0.05)
Teacher experience				-0.02	-0.04**	-0.02
				(0.02)	(0.02)	(0.03)
Teacher gender				0.02	-0.00	0.01
				(0.09)	(0.11)	(0.12)
Constant	-0.53***	-0.59**	-0.77***	-0.71**	-0.53	-1.13***
	(0.19)	(0.28)	(0.22)	(0.28)	(0.37)	(0.37)
Observations	2,692	1,276	1,416	2,656	1,261	1,395
R-squared	0.07	0.04	0.06	0.10	0.06	0.10

Note: All estimations included dummy y2016, y2016*male and the square of teacher experience. The corresponding coefficients are insignificant and have not been reported. Robust standard errors in parentheses p<0.01, ** p<0.05, * p<0.1

Table 5: Oaxaca-Blinder decomposition of differences in math scores between boys and girls

	Threefold decomposition		Twofold decomposition (weight=0.5)	
	With school fixed effects[1]	No school fixed effects[2]	With school fixed effects[1]	No school fixed effects[2]
Prediction females	0.20***	0.20***	0.20***	0.20*
	(0.03)	(0.06)	(0.03)	(0.06)
Prediction males	-0.01	-0.01	-0.01***	-0.01
	(0.02)	(0.06)	(0.02)	(0.06)
Difference	0.21***	0.21***	0.21***	0.21***
	(0.02)	(0.05)	(0.02)	(0.04)
Decomposition				
Endowments	-0.00	-0.02	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)
Coefficients	0.25***	0.24***	0.23***	0.23***
	(0.01)	(0.04)	(0.01)	(0.04)
Interaction	-0.04	-0.01		
	(0.01)**	(0.02)		
Observations	2656	2656	2656	2656

Note: [1]Estimation uses variables in specification 2 with standard errors accounting for fixed regressors (i.e., the school dummies). [2]Estimation excludes school dummies but includes principal gender and professional grade, school type. Robust standard errors adjusted for school clusters in parentheses*** p<0.01, ** p<0.05, * p<0.10

Table 6: Oaxaca Decomposition of standardised maths scores between boys and girls using a reweighted recentered influence function model

	quantile									
	10		30		50		70		90	
Overall										
Prediction females	-1.03***	(0.05)	-0.48***	(0.07)	0.04	(0.08)	0.74***	(0.09)	1.71***	(0.10)
Prediction males	-1.15***	(0.04)	-0.72***	(0.06)	-0.22**	(0.08)	0.51***	(0.11)	1.52***	(0.12)
difference	0.12**	(0.05)	0.25***	(0.05)	0.26***	(0.06)	0.23***	(0.08)	0.20***	(0.07)
Explained	-0.01	(0.01)	-0.01	(0.01)	-0.01	(0.02)	-0.00	(0.02)	-0.06**	(0.03)
Unexplained	0.12**	(0.05)	0.26***	(0.06)	0.28***	(0.06)	0.23***	(0.08)	0.26***	(0.07)
Explained										
Mosteduc	-0.00*	(0.00)	-0.01***	(0.00)	-0.02***	(0.00)	-0.02***	(0.01)	-0.02***	(0.01)
Aspirations-university	0.00	(0.00)	0.01**	(0.00)	0.02***	(0.01)	0.03***	(0.01)	-0.00	(0.01)
Teacher rating	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.01**	(0.00)	0.00	(0.00)
Unexplained										
Aspirations-university	0.06	(0.08)	-0.08	(0.05)	-0.03	(0.07)	-0.01	(0.09)	0.26***	(0.09)
Teacher rating	0.60**	(0.24)	0.93***	(0.39)	0.70*	(0.38)	0.00	(0.45)	0.49	(0.51)

Note: Explanatory variables also included household income, ethnicity, teacher gender, teacher experience, year, psi*year and school fixed effects. Coefficients not been reported as they are not statistically significant apart from the ethnicity variable Tamil at the 10^{th} percentile and Sinhala at the 90^{th} percentile. Standard errors in parenthesis *** $p<0.01$, ** $p<0.05$, * $p<0.10$. Observations: 1395 female, 1261 male.

Table 7: Mean characteristics of 15-25 year olds in the sample

	Females	Males	Difference in means
Enrolment in full time education			
Age 15-17	0.86	0.85	0.01(0.01)
Age 18-19	0.55	0.53	0.02(0.02)
Age 20-23	0.19	0.16	0.03**(0.05)
Age 24-25	0.08	0.04	0.04***(0.01)
Completed years of education			
Age 15-17	9.83	9.73	0.10**(0.04)
Age 18-19	10.82	10.51	0.31***(0.08)
Age 20-23	10.97	10.45	0.51***(0.07)
Age 24-25	11.01	10.21	0.80***(0.11)
Unemployment rate[1]	0.38	0.27	0.11***(0.01)
Labour force participation[2]	0.27	0.54	-0.27***(0.01)
Wage earnings (rupees)[3]	12,029.50	14,099.57	-2010.07***(358.72)
All earnings (rupees)[4]	12,559.26	15,519.55	-2960.29***(528.29)
Predicted wage	4900.85	6221.84	-1320.99***(54.19)
Predicted earnings	5250.54	8019.46	-2768.92***(65.05)
Urban residence (proportion)	0.26	0.26	-0.00 (0.00)
Sinhala (proportion)	0.62	0.61	0.01 (0.01)
Tamil (proportion)	0.25	0.25	0.00 (0.01)
Wealth index	0.55	0.55	0.00(0.00)
Household size	4.82	4.83	-0.02(0.03)
Observations	6727	6018	

Notes: [1]unemployed/(employed+unemployed). [2](employed+unemployed)/total. [3] Based on positive wage earners only: 935 women and 1757 men. [4] Based on positive earners: 1010 women and 2090 men. Standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.10.

Table 8a: What influences enrolment beyond compulsory years among 15-25 year-olds? Average Marginal Effects

Dependent variable: enrolled=1, 0 otherwise

	Full sample		Subsample: Those who've completed secondary education	
	(1)	(2)	(3)	(4)
male	-0.05***	-0.04***	-0.04	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Age 17-18	-0.26***	-0.25***		
	(0.01)	(0.01)		
19-20	-0.56***	-0.55***		
	(0.01)	(0.02)		
21-23	-0.67***	-0.66***	-0.10*	-0.09
	(0.01)	(0.01)	(0.06)	(0.06)
Wealth index	0.56***	0.56***	0.64***	0.63***
	(0.03)	(0.03)	(0.13)	(0.13
Opportunity cost	-0.01***	-0.01***	-0.02**	-0.02***
	(0.00)	(0.00)	(0.01)	(0.01)
Teacher	0.14**	0.14**		
experience	(0.05)	(0.05)		
School IAC	-0.34**	-0.35**		
	(0.15)	(0.15)		
Never married	0.31***	0.31***	0.37***	0.37***
	(0.01)	(0.01)	(0.01)	(0.01)
hhsize	-0.01***	-0.01***	-0.01	-0.01

	(0.00)	(0.00)	(0.00)	(0.01)
ethnic_Sinhala	0.07***	0.07***	0.0	0.04
	(0.01)	(0.01)	(0.03)	(0.03)
observations	12740	12740	1687	1687

Note: All estimations include province level dummies, urban dummy and Tamil ethnicity dummy. Coefficients not reported. Standard errors (delta method) reported in brackets. ***p<0.01, ** p<0.05, * p<0.1. The corresponding probit results for specification 1 and 2 can be found in Appendix Table A1.

Table 8b: Pairwise comparisons of average marginal effects for males versus females

	Contrast dy/dx	Standard error (delta method)	Unadjusted [95% Conf. Interval]	
Specification 1 (opportunity cost measured using wages)				
Wealth	0.12	0.05	0.02	0.22
Opportunity cost	-0.01	0	-0.02	-0.01
Teacher experience	0.05	0.08	-0.11	0.2
School IAB	0.53	0.22	0.11	0.96
Specification 2 (opportunity cost measured using total earnings)				
Wealth	0.12	0.05	0.02	0.22
Opportunity cost	0.01	0	-0.02	-0.01
Teacher experience	0.07	0.08	-0.09	0.22
School IAB	0.53	0.22	0.11	0.96

The corresponding probit results for specification 1 and 2 can be found in Appendix Table A1, column 1 and 2.

.

Table 9: Average education levels and earnings by occupation category for those aged 15-30 in 2012/13

Occupational categories	Earnings in rupees			Education in years			Women/(all employed in category)	Women/(all employed women)	Men/(all employed men)
	Women	Men	Diff. (1)-(2)	Women	Men	Diff. (4)-(5)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Managers	21928.8	40649.4	-18720.6 (14041.2))	11.60	11.10	0.5*(0.2)	0.33	0.04	0.04
Professionals	17884.3	27805.1	-9920.8*** 1723.8	13.06	12.84	0.22(0.18)	0.67	0.13	0.03
Technicians	17366.9	23188.8	-5821.9*** (1511.7)	13.09	11.94	1.15***(0.1)	0.39	0.12	0.09
Clerical support	16946.2	18959.2	-2012.9** (1006.0)	12.78	12.23	0.55***(0.1)	0.62	0.13	0.04
Sales	12347.9	15888.4	-3540.6** (1529.5)	11.18	10.43	0.75***(0.1)	0.28	0.09	0.11
Skilled agri/fish	3069.8	9389.3	-6319.5***(1202.0)	10.02	9.28	0.74***(0.2)	0.20	0.06	0.12
Skilled craft	9618.77	16849.9	-7231.2*** (924.6)	10.16	9.62	0.54***(0.1)	0.32	0.18	0.18
Plant/mach. Op.	11868.1	18436.8	-6568.73**(2721.20)	10.22	9.73	0.49**(0.2)	0.17	0.05	0.12
Elementary	9282.7	12850.7	-3567.9***(609.5)	7.95	8.43	-0.48**(0.2)	0.25	0.19	0.27
Overall	12971.7	16805.5	-3878.7***(711.5)	10.58	10.08	0.49***(0.0)	0.32	1.00	1.00

Values based on 2131 women and 4543 men. Standard errors reported in brackets *** p<0.01, ** p<0.05, * p<0.1

Table 10: Labour force participation rates and marital status among 30-49 years olds

		Did not complete secondary school (i.e., did not pass Advanced Level or equivalent)	Completed Secondary School (i.e., passed Advanced Level exams or equivalent)	Educated beyond secondary school (i.e., Advanced Level+vocational or professional qualification, university degree)	All education levels
Labour force participation (%)[1]	Female	37.8	48.7	92.7	42.0
	Male	97.0	98.2	99.4	97.3
Median age at first marriage[2]	Female	20.7-23.6	26.2	28.7	23.3
	Male				
Never married(%)[3]	Female	5.2	7.9	11.9	5.9
	Male	8.1	9.7	9.4	8.4

[1] Calculated as (employed+unemployed)/total in relevant education group, based on HIES 2012 data.
[2] Taken from DHS, (2017, p.83). [3.] Based on HIES 2012 data.

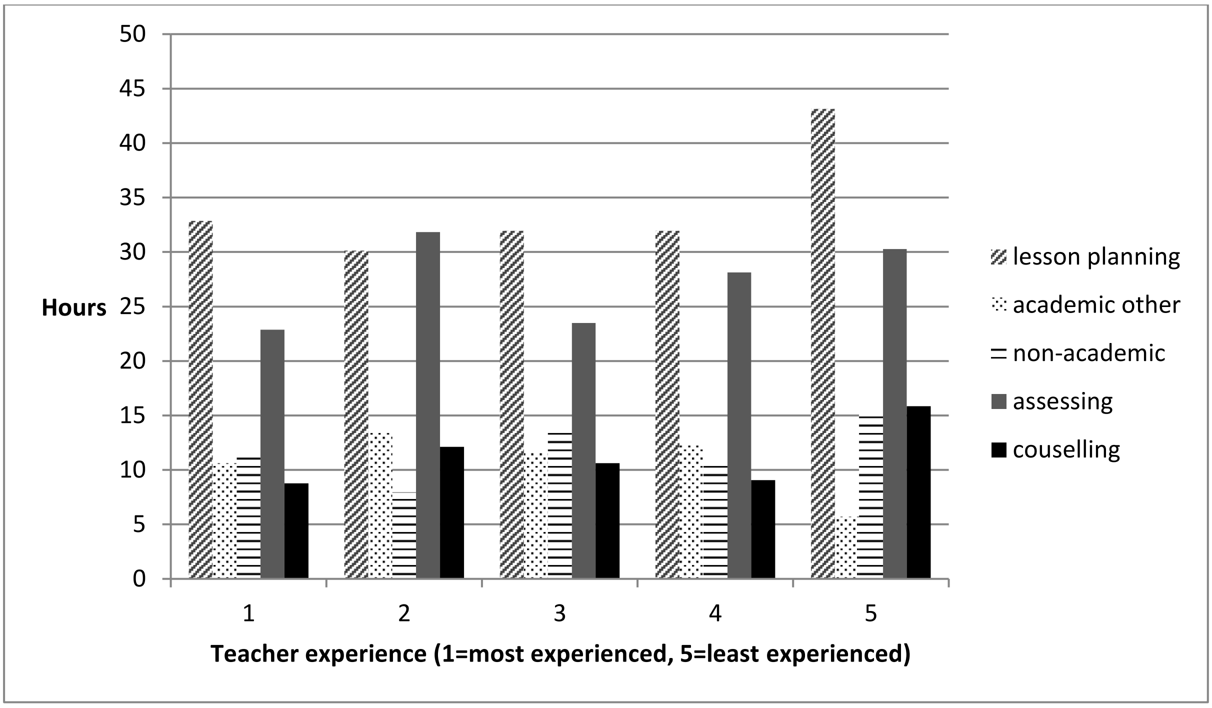


Figure 1: Teacher time-use on selected aspects during the preceding term

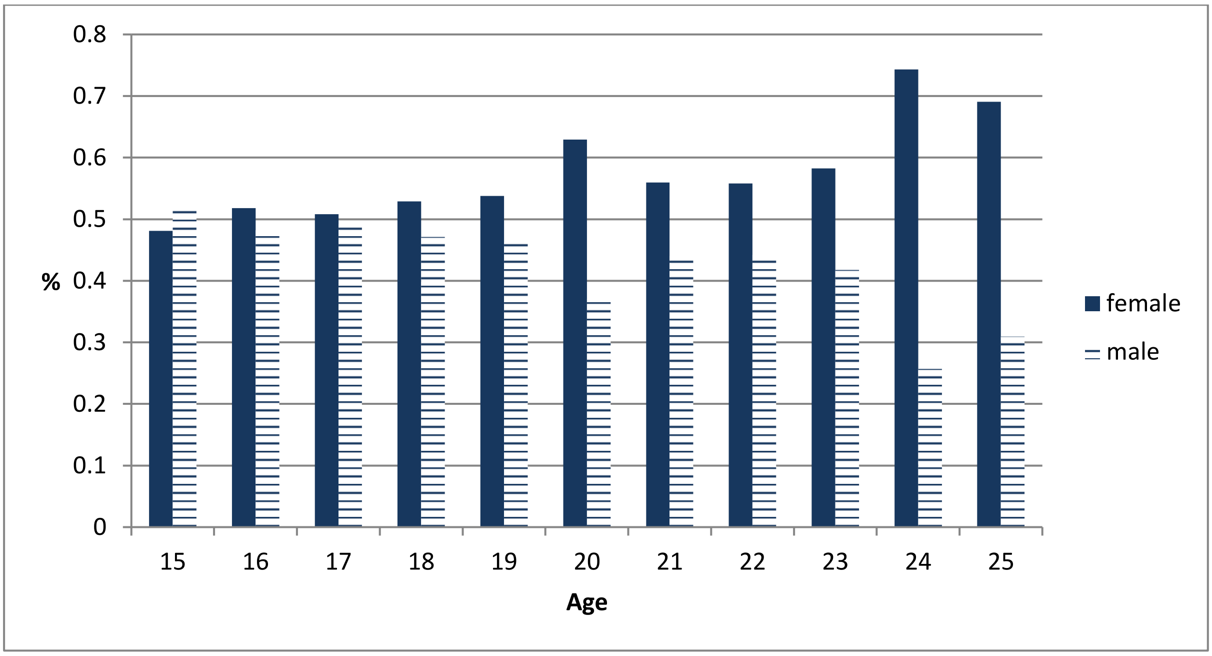


Figure 2: Gender composition of students enrolled in full-time education

Notes: Compulsory schooling is only up to age 14. Thus the horizontal axis represents education beyond the age of compulsory schooling.

Appendix

Figure A1 Net enrolment in Secondary Education

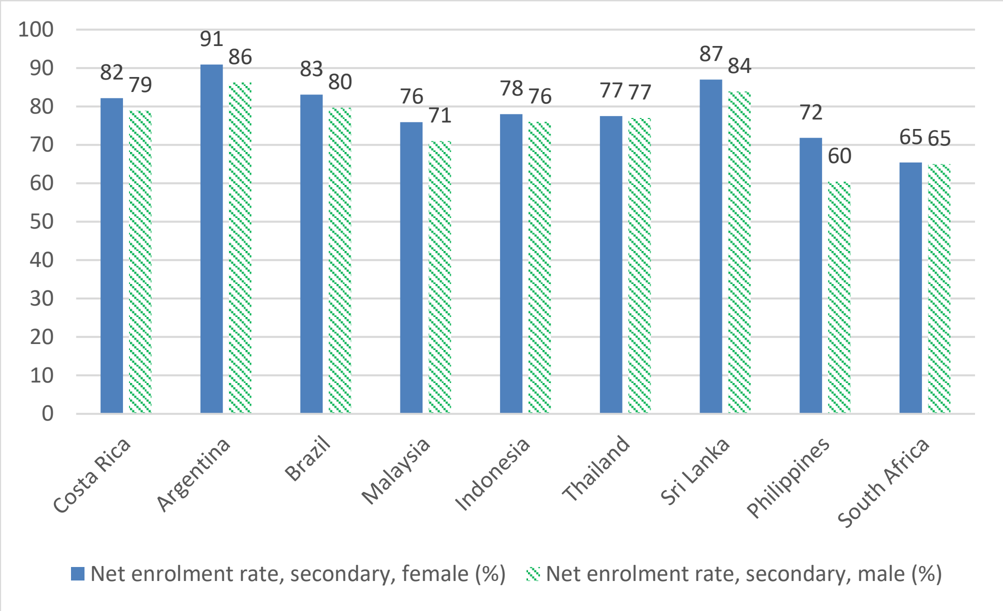


Figure A2 Net enrolment in Tertiary education

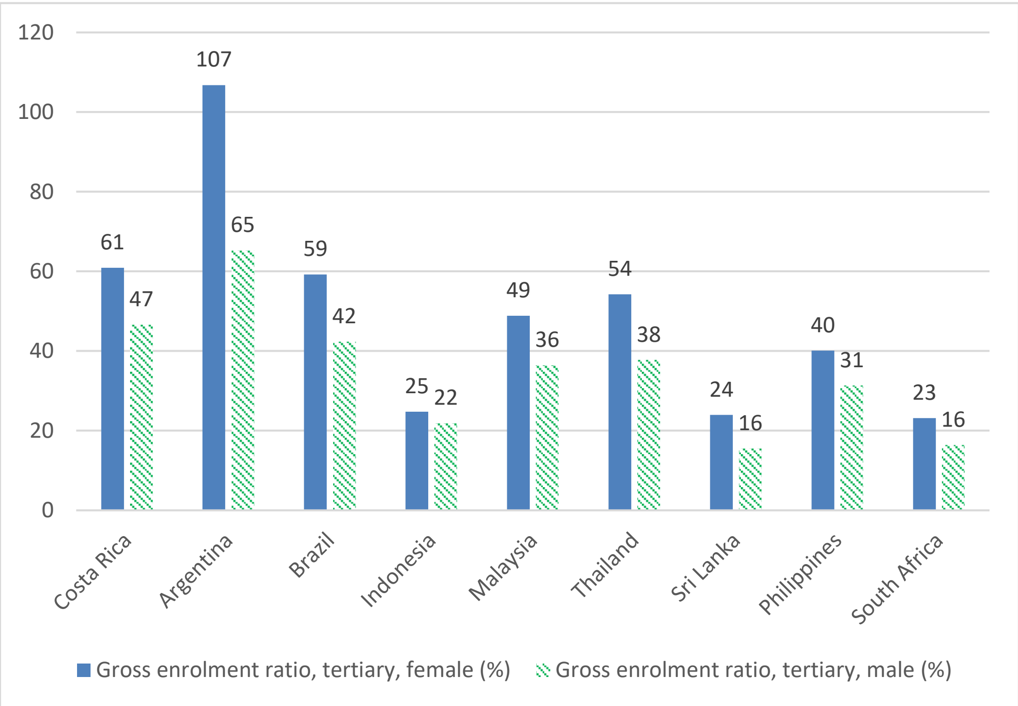
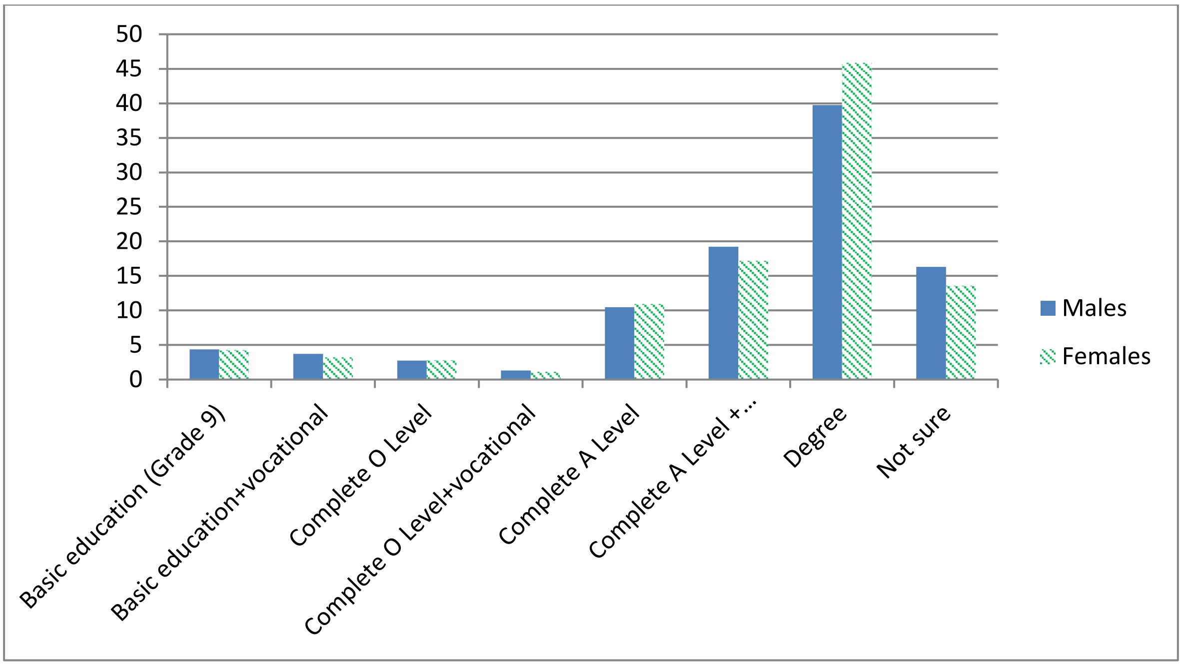


Figure A3. Parental/child aspirations for child's education



Response to question 'What is your child's plan after Compulsory (Basic) Education?' asked from child's parent or guardian in 2012, when child was in grade 7 (age 12).

Figure A4: Occupational variation by men and women between ages 15-25.

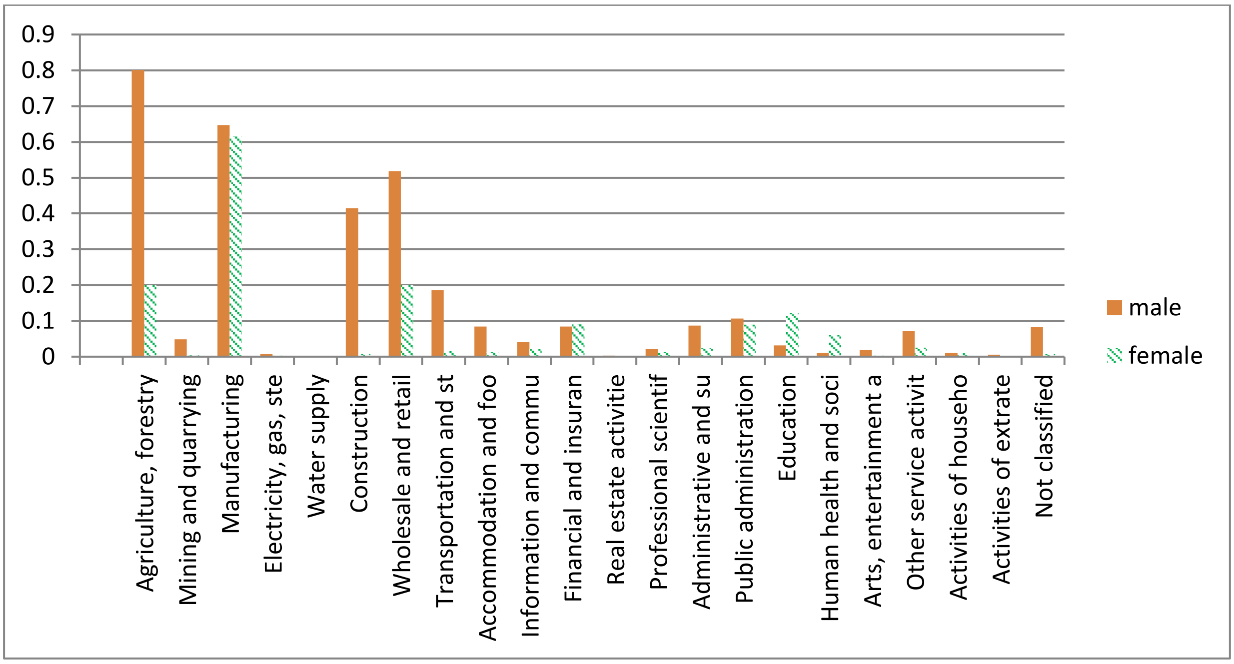


Table A1 What influences enrolment beyond compulsory years among 15-25 year-olds?

Deependent variable =1 if enrolled, 0 otherwise

	Full sample		Subsample: Those how have completed secondary education (i.e., completed Advanced Level, grade 13)	
	(1)	(2)	(3)	(4)
Age 18-19	-0.89***	-0.87***		
	(0.04)	(0.04)		
Age 20-23	-1.90***	-1.88***		
	(0.04)	(0.05)		
Age 24-25	-2.46***	-2.42***	-0.36*	-0.31
	(0.08)	(0.08)	(0.20)	(0.22)
wealth	2.21***	2.19***	1.49***	1.51***
	(0.17)	(0.17)	(0.58)	(0.57)
male	-0.68***	-0.70***	-0.91*	-0.80
	(0.20)	(0.21)	(0.49)	(0.52)
male#wealth	0.72***	0.75***	1.70**	1.65*
	(0.24)	(0.24)	(0.86)	(0.86)
oppcost	0.00	0.00	-0.07***	-0.07***
	(0.01)	(0.01)	(0.03)	(0.03)
male#c.oppcost	-0.06***	-0.05***	-0.03	-0.02

teacher_exp	(0.01)	(0.01)	(0.04)	(0.03)
	0.54*	0.49		
male#c.teacher_exp	(0.30)	(0.30)		
	0.25	0.35		
school_IAB	(0.36)	(0.36)		
	-2.67***	-2.68***		
male#c.school_IAB	(0.86)	(0.86)		
	2.29**	2.27**		
nevermarried	(0.97)	(0.97)		
	(0.06)	(0.06)	(0.16)	(0.16)
hhsize	-0.03***	-0.03***	-0.04	-0.04
	(0.01)	(0.01)	(0.03)	(0.03)
urban	0.04	0.03	-0.09	-0.11**
	(0.04)	(0.04)	(0.06)	(0.05)
Sinhala	0.32***	0.32***	0.12	0.12
	(0.05)	(0.05)	(0.12)	(0.12)
Tamil	-0.04	-0.05	0.08	0.08
	(0.06)	(0.06)	(0.16)	(0.16)
Constant	-1.50***	-1.48***	-1.46***	-1.43***
	(0.22)	(0.22)	(0.40)	(0.40)
Observations	12,740	12,740	1,687	1,687

Notes: Probit estimation output. Province included in estimation but results not reported.

Standard errors in parentheses*** $p<0.01$, ** $p<0.05$, * $p<0.1$