SHIV NADAR UNIVERSITY

DEPARTMENT OF ECONOMICS
WORKING PAPER SERIES

Working paper # SNUECON-WP2016-002

Dakar, are we there yet ! Analysing Gender Gaps in Learning in Rural India

CHANDAN JAIN

http:/www.snu.edu.in/papers/wp2016002

Department of Economics
P.O. Shiv Nadar University
Greater Noida
Uttar Pradesh – 201314, India.

October 2016

Dakar, are we there yet !
Analysing Gender Gaps in Learning in Rural India

Chandan Jain *

Shiv Nadar University

October 6, 2016

Abstract

In the present study, we use three rounds (2009, 2012 and 2014) of ASER data on learning outcomes from rural India in order to analyse gender gaps in performance on standardized tests of reading, maths and english reading ability. Despite the existence of a large body of literature analysing gender gaps in learning in the international context, the study of the same remains limited in the Indian context. Further, our data also allows us to compare these gaps for the pre and post Right to Education Act (RTE) period as well. Using siblings fixed effects estimation we find that, at the baseline level, boys outperformed girls across all the three tests for the 2009 round of survey and that these gaps have widened overtime in the post RTE period, except for the case of reading wherein girls seem to be performing better overtime. Additionally, we find that mother's education tends to close in on the above gap and that girls belonging to educated mother tends to perform better in comparison to boys, while private school enrolment and tuition access tends to widen the gap as a result of a differential effect of the same on learning across gender. We attribute the latter effect to differences in expenditure across gender, using data for expenditure on tuitions, we find that on average the expenditure on tuitions for girls tends to be lower as compared to the same for boys. We also find that the gaps in learning tends to be lower across states that are more gender equal. Lastly, using data on enrolment to private school and tuition as proxies for household division of resources, we examine the change in the same overtime. We find that overtime the distribution of resources has moved more in favour of boys than for girls.

* PhD Student in Economics at Shiv Nadar University (SNU). I would like to thank Shampa Bhattacharjee, Ashokankur Datta and Gitanjali Sen for their suggestions on the same. The data used in the paper was shared by the ASER Centre as a part of the ASER Call for Paper 2015. Mail: cj656@snu.edu.in.

Introduction

In the year 2000, a total of around 164 countries adopted the Dakar Framework for Action, within the same these countries committed themselves to providing quality education to their citizens and achieving the goal of Education for All by the year 2015. The framework included a total of six goals aimed at improving educational outcomes and within the same a major emphasis has been on reducing gender disparities in educational access, attainment and learning outcomes. In pursuit of the same, various countries overtime have implemented programmes and policies towards reducing gender disparities in education, few of these includes - Female Stipend Programme in Bangladesh, Girls Scholarship Programme in Kenya, Bicycle Programme in Bihar, India. While reducing gender disparities in education access and attainment might seem justified on grounds of social justice, the same also becomes relevant from an economics perspective as well. The reason for the same being, gender disparities in education attainment as well learning outcomes can have severe implications at both the micro as well as the macro level. The negative link between gender inequality in education and economic development has been documented by a number of studies, both theoretical as well as empirical, few such studies include Klasen (2002), Klasen & Lamanna (2008). Similarly, Schultz (2002) notes that the marginal social returns from women's level of education are much larger in comparison to the same from men's levels of education. At the individual level on the other hand, gender differences in the level of educational attainment, learning and related abilities (such as maths and reading skills) could further substantiate differences in labour market outcomes across the two groups. Given this understanding of the negative effects of gender gap in education and learning on economic development as well as on individual level welfare, the present paper analyses gender gaps in learning outcomes in the context of rural India [1]. Despite this being an active area of research in the international context, the study of gender gaps in learning remains under studied in the Indian context. The present study, therefore, aims to fill in this gap.

In India, boys have remained as the more favoured group for a large part of the country and girls have been discriminated against with regard to household division of resources, education access as well as attainment. The same in turn has resulted in low education attainment for girls, especially in the more backward and rural regions of the country. An instance for the same can be seen from the low literacy rate for females in the country, as per the Census 2001 data only 53.67% of females in India were literate while for men the same was around 75.26%. In order to close in on such disparities, starting 2003-04 and alongside its flagship education programme (Sarva Sikhsha Abhiyan), Government of India implemented two programmes (NPEGEL and KGBV)[2] which aimed at raising educational attainment and learning outcomes for girls in the country. A further push in this direction has been the Right to Education Act, passed in the year 2009 the same provides for free education to all until 14 years of age. These programmes and the related Sarva Sikhsa Abhiyan have been quite successful in raising enrolments at the elementary level for both boys as well as for girls. Concerns, however, remain with regard to learning outcomes. Evidence based on the recent ASER surveys show that the learning outcomes in India, measured using performance on standardized tests for maths and reading, have been on a decline despite the various educational reforms undertaken in the past one decade. Given this decline in learning outcomes, what remains unknown is - are there differences between boys and girls with regard to their learning and how has the same changed overtime. The present paper aims to answer the same, we analyse the following questions in this paper - i) how significant is the gap between boys and girls with regard to their learning outcomes, wherein learning outcomes have been measured using performance on standardized tests for maths, reading and english reading ability in rural India, ii) how has the same changed overtime in the post RTE period, iii) what are the different factors that contribute to the observed gender gaps in learning, and iv) how has the distribution of household resources between boys and girls changed over time. We analyse the same using data on learning outcomes from three rounds of ASER survey [3] and our pooled sample

[1] The same comprises of around 70 % of the entire population in India

[2] NPEGEL refers to National Programme for Education of Girls at the Elementary Level, while KGBV refers to Kasturba Gandhi Ballika Vidhalaya Scheme

[3] We limit our attention to rural India only because of the lack of similar data for the urban parts of the country.

comprises of around half a million observations for children in the age group 10 to 16 years, who were tested for their reading, mathematical and english reading ability within the 2009, 2012 and 2014 round of ASER survey. The 2009 round helps us to explore the situation for the pre RTE period while the later two rounds belong to the post RTE period. Using siblings fixed effects estimation we find the following - a) for the 2009 round of survey, girls, at the baseline [4], had a lower probability of scoring the highest level for all the three assessment tests conducted within the survey, and b) the same has further increased over time for the 2012 and 2014 round of survey, except for the case of reading assessment wherein some convergence can be seen as girls seem to be performing better than boys overtime. For the factors that affect the observed gender gap, we find that an increase in the mothers level of education tends to narrow down the above observed gap between boys and girls and girls belonging to educated mother tends to perform better in comparison to boys. Additionally, we find that, although private school attendance and tuition access benefits both boys and girls but the girls seem to gain less in comparison to the same for boys. To dwell into the reason for this observed differential effect, we use data on the amount spent on tuition for each child from the 2014 round of ASER survey, using a siblings fixed effects estimation we find that the amount spent on tuitions is lower for girls as compared to the same for boys, thus suggesting that even though boys and girls might have access to the same resources but there still can be discrimination with regard to their quality. Further, in order to analyse the role of culture and attitudes towards women in explaining the observed gender differences in learning, we divide our sample into states that are more gender equal and those that are more unequal, using the same we find that for states that are more gender equal the gaps in learning tends to be lower in comparison to states that are more unequal. Lastly, using data on access to private schools and tuitions from the ASER survey we examine the changes in household division of resources between boys and girls overtime. Again, based on siblings fixed effects estimation we find that overtime the division of household resources has tended to move in favour of boys.

A study closest to ours is a study by White, Ruther & Kahn (2016), which uses the IHDS 2004-05 round data in order to examine the different factors that affects the gender gap in reading, writing and maths performance for children in 8 to 11 years of age in India. Compared to White, Ruther & Kahn (2016), we not only look at the existence of gender gaps in learning for a particular year but also at the changes in the same overtime for the period starting 2009. The same also allows us to capture the changes in learning for the post education reform period, as the majority of educational reforms such as SSA, NPEGEL and KGBV have been undertaken around the 2004-05 period, while the study by White, Ruther & Kahn (2016) is limited to the analysis of gender gaps for a single time period only, which is 2004-05. Similarly, the 2012 as well as the 2014 round of the survey allows us to look at existence of the gender gaps in the post RTE period as well. Additionally, compared to White, Ruther & Kahn (2016) our sample relates to children in 10 to 16 years of age group and our entire pooled sample consists of around half a million observations restricted only to rural parts of the country. White, Ruther & Kahn (2016) on the other hand uses a sample of around less than 15,000 observations, which is drawn from both urban as well as rural parts of the country. Finally, our data also allows us to look at the gender gaps with respect to english reading ability as well, the reason that we focus on english reading ability is that given globalization and interconnected markets the same might be associated with positive returns in labour markets. Different studies (for instance Bleakly & Chin (2003), Lang & Siniver (2006)) in the past have shown that language skills (particularly english speaking ability) are associated with benefits in the labour market. For the case of India as well, Azam, Chin & Prakash (2011) have shown that the ability to speak english is associated with an increase in the hourly wages. Similarly, the positive link between maths ability and gains in labour market outcomes have been well documented, few studies in this regard includes Rose & Betts (2004), Arcidiacono (2004), Joenson & Nielson (2006).

The above evidence for the gains from english and maths ability in labour market forms a compelling

[4] The baseline is defined as children belonging to mother's with zero years of education and are also not enrolled in private schools or tuitions.

reason for our focus on gender gaps in learning outcomes rather than focussing on mere educational attainments as measured using the specific grade completed. Learning outcomes, in the present study, measured using performance on standardized tests for reading, maths and english reading allows us to capture differences between individuals with regard to their quality of learning, which is not possible using measures on educational attainment. The same becomes much more important in the Indian context for the reason that there could be discrimination between boys and girls with regard to the type and quality of schools or resources that they are provided with. We provide evidence for the same using data on expenditure incurred on tuitions for each child. Data on test scores allows us to capture for such discrimination. In the recent literature on links between education and economic development (see Hanushek & Woessman (2008)) as well, emphasis has been on using cognitive skills that allows for a measure of quality differences rather than mere education attainment as determinants of income levels as well as economic growth.

Finally, our focus, herein on analysing gender gaps in the pre and the post RTE period is for the reason that on one hand the implementation of the RTE Act represented a positive move towards improving educational outcomes by making the government responsible for providing free as well compulsory education to all children in the 6 to 14 year age group; while on the other hand a specific provision under the RTE act might possibly have a negative impact on learning outcomes. The same being the *No Detention Policy* under Section 30 (1) of the RTE Act, as per which no child can be detained or held back in a grade till the completion of elementary education. In the same regard, a number of state governments (such as Delhi, Rajasthan) in the recent past have requested the central government to review the no detention policy under the RTE Act. Given this, our interest therefore lies in assessing gender gaps in light of these two differential effects that the RTE Act could have on learning.

The entire study has been divided into different sections. Section 2 presents a review of the past studies that have analysed gender gaps in the performance on standardized tests of reading and maths ability. Section 3 describes the data that we use for our analyses and also our working sample. Section 4 discusses the conceptual framework as well the details the methodology employed within the paper. Section 5 discusses the dependent variable and presents a descriptive analysis. Section 5 presents the main results for the paper while section 6 concludes the study.

Literature Review

There exists a vast body of literature analysing gender differences in performance on standardized tests of reading, writing, mathematical and science ability. Despite the existence of a vast body of work in this field, evidence in the context of India remains sparse. Only a recent paper by White, Ruther & Kahn (2016) has attempted to analyse these gender differences and their causes in the context of India. As already discussed above, the study by White, Ruther & Kahn (2016) uses the IHDS 2004-05 round of data in order look at factors which could explain the gender gap in performance on reading and maths tests for children between 8 to 11 years of age, it uses an ordered logit regression framework in order to analyse the impact of social background, access to learning resources, cultural attitudes and time allocation on the same. For the various factors analysed, White, Ruther & Kahn (2016) finds that girls living with younger siblings tends to have a lower score in both reading as well as mathematics and additionally the economic status of the household measured using the household asset index contributes towards minimizing the observed gap; while it does not find the effect of any other factor analysed to be statistically significant in explaining the gender gap being studied.

Given above, studies done in the international context have focused either on documenting the gap that exists (for instance Xin (2008)[5]) or have tried explaining the same for different regions, sub regions and age groups, with a larger focus on explaining gender gaps specifically with regard to performance on maths based tests. The reasoning for the same, as provided by these studies, being

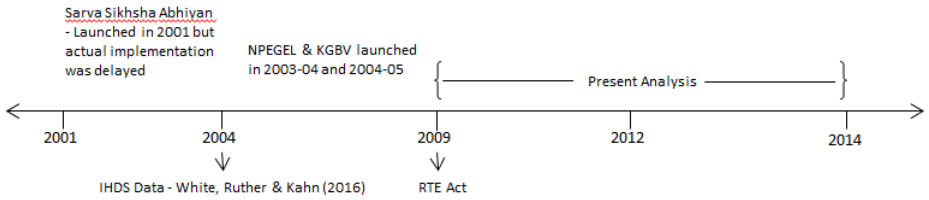
[5] Xin (2008) for instance compares the mean difference in the performance of boys and girls over different standardized tests for learning conducted under different programmes such PISA, TIMSS, PIRLS, PASEC etc.

that mathematical capabilities of an individual are better at predicting future incomes as well labour market outcomes for individuals in the long run. A common result that follows from such studies is that boys on average tends to perform better than girls with regard to maths based tests while it's the other way round for the case of language and reading based tests. With regard to what explains these gaps, the role of a number of factors have been analysed in the past literature. Guiso (2008), for instance, attributes the gender gap in performance on standardized tests for maths to differences in the status of women in the society, wherein it measures the status using World Bank Gender Gap Index. Using PISA 2003 data for scores on maths test across countries, Guiso (2008) finds a positive correlation between Gender Gap Index and the gap on performance in maths tests. Building on Guiso (2008), Nollenberger, Planas & Sevilla (2014) have tried to disentangle the role of culture from that of the institutions in trying to explain the gender gap in maths tests from PISA, to do so it relies on second generation migrants and it finds that a higher degree of gender equality in the country of ancestry is associated with a lower gender gap in maths tests for the second generation migrants. Another study examining the role of gender equality in the society is the study by Pope & Syndor (2008), using data from US states it finds that gender gaps on reading and maths tests tends to higher in states that are more gender unequal. Other than gender equality, study by Fryer & Levitt (2009) have looked into the role of socio economic conditions in explaining the observed gender gaps in reading and maths tests for children in the US, however, it finds no substantial effect of the same. In addition, Fryer & Levitt (2009) also undertakes a cross country comparison in order to explore the findings by Guiso (2008) using the TIMSS dataset. Contrary to Guiso (2008) , Fryer & Levitt (2009) finds no effect of the Gender Gap Index on the observed gender gap for maths tests, with this non significant effect being driven by countries that are highly unequal in terms of gender but these still have lower lower gender gaps in maths tests. Fryer & Levitt (2009) contributes the same to presence of single sex schools in such countries. Towards the same, study by Bedard & Cho (2009) and Doris, Neill & Sweetman (2012) have looked into the impact of school related factors on gender differences in learning. Bedard & Cho (2009) have examined the role of academic sorting and gender composition of class for the case of OECD countries while Doris, Neill & Sweetman (2012) have examined the impact of single sex schools for the case of Ireland. Bedard & Cho (2009) finds that academic sorting as well as the gender composition does matter for the observed gender gaps in learning while Doris, Neill & Sweetman (2012) does not find any impact of single sex schools on the same.

Indian Context - A Decade of Educational Reforms

For the period starting 2000-01, Government of India overtime has implemented a series of educational reforms in the past one decade. At the same time in 2001, the gross enrolment ratio (GER) at the elementary level in India was around 81.6, with boys having a higher GER of around 90.3 as compared to 72.4 for girls (MHRD 2014). Similarly, for the case of literacy rates as well there existed a gap of close to around 20 percentage points between men and women for the year 2001 (Census 2001). Towards the same, government implemented the Sarva Sikhsha Abhiyan (SSA), in the year 2001, with the aim to universalize elementary education in the country. Implementation of the SSA has been followed by the subsequent implementation of the Midday Meal Scheme (MDM), which required government primary schools to provide mid day meals to children. Additionally, SSA was later followed by two additional schemes (NPEGEL & KGBV), implemented in 2003-04 and 2004-05, that particularly targeted at improving educational outcomes for girls in the elementary school going age. Both the NPEGEL and KGBV programmes have been supply side interventions aimed at improving school access as well the existing school infrastructure as per the girls' requirements. A following intervention has been the Central Education Institutions Act 2006, which provides for the reservation of OBCs across higher educational institutions in the country. Finally, the last major reform has been the Right to Education Act 2009 and the same provides for free and compulsory education to everyone in 6-14 year age group in India. RTE Act binds the government to ensure admission, attendance as well completion of education for all children in 6-14 years of age group. The figure below gives a timeline for the above discussed interventions in India and presents the time period that present study focusses on in the analysis below and the same also presents the period analysed by the previous study

by White, Ruther & Kahn (2016) as well.



As already stated before in the introductory section, the above discussed reforms have been quite successful in raising enrolment at the elementary level for both boys as well as for girls, for boys the GER was around 104.5 at the elementary level for the year 2010-11 while for girls the same was around 103.3 (MHRD 2014). The thing that is however disturbing is that, despite these educational reforms as well increases in educational participation at the elementary level, learning outcomes in India (particularly rural India) have been on a decline in the recent time. From the recent ASER reports one can see that learning outcomes in rural India have seen decline overtime. For instance, the percentage of children, enrolled in class 8th, who can read a standard 2 level text has declined overtime between the ASER 2008 and 2014 round of survey (ASER 2014). Similar results can be observed in the case of mathematics as well wherein overtime there has been a decline in children ability to do basic mathematics. A similar observation can be seen from table 2 as well, wherein we discuss in detail the evolution of learning outcomes across the different ASER rounds considered in the analysis below.

Data

The present analysis uses data from three different rounds (2009, 2012 and 2014) of the ASER Survey. ASER survey is a nationally representative survey of learning and is conducted annually across a majority of districts in the country. The survey, however, is restricted to rural areas of the district only. Within each district a total of of around 30 villages are selected for survey and within each village a total of around 20 households are selected. These households are selected in a way to ensure a maximum possible spatial coverage across the village chosen for the survey. Within the selected household all children between age 5-16 are tested for their learning abilities using standardized tests on reading, maths and english reading and for each child his/her performance on these tests is recorded. In addition to these learning tests, for each of the household selected for survey, ASER also collects data on a number of indicators relating to children (such as their age, gender, school enrolment, school type, tuitions etc.), their parents (age and education) and different household indicators (such as availability of electricity and toilet in the household). Besides, for each of the village selected for sampling ASER also collects data on different village level indicators for access to different services such as availability of primary, middle, secondary school, medical facilities, electricity access etc. Finally, the reason we restrict ourselves to these three rounds in specific is that unlike the other rounds of ASER that tests children only for their reading and maths ability, these three rounds tested children on their english reading ability as well and thus provides us with an additional dimension for learning. For reasons explained ahead in the text, we focus only on children who are aged 10 or above. Given that we would be using siblings fixed effects estimation in our analysis ahead, therefore, we further limit our sample to only those mothers with 2 or more children. We also drop all missing observations from our sample and accordingly our working sample consists of a total of around 5,29,078 children of 10-16 years of age, who completed all the three assessment tests within the ASER survey. The same amounts to an average of around 0.17 million observations for each round under consideration. All our results in the following sections below are thus based on the working sample only.

Learning Outcomes in ASER Surveys

Within the ASER survey each child in the household is tested on three different assessment tests, comprising of a test on reading, maths and english reading ability and based on their performance

they are accordingly ranked. Column 2 in table 1 describes the various levels that a child can achieve, for instance if a child is unable to read then he/she is assigned to the category *Read Nothing*.

Table 1: Learning Assessment in ASER

Assessment	Assessment Criteria
Reading Assessment (R)	Read Nothing Can Read Letters Can Read Words Can Read Paragraphs Can Read Stories
Maths Assessment (M)	Nothing Can Recognize Numbers (1-9) Can Recognize Numbers (10-99) Subtraction Division
English Reading Assessment (E)	Read Nothing Can Read Capital Letters Can Read Small Letters Can Read Words Can Read Sentences

Conceptual Framework and Methodology

As described before, our aim in this paper is to analyse the existence as well as the evolution of gender gap overtime. A simple approach to accomplish the same could be to use descriptive statistics and analyse the changes in the same overtime, although useful, the same however fails to answer as to whether the observed gaps are statistically significant or not when different factors such as parents education or income levels are controlled for. To overcome the same, herein, we analyse gender gap using a siblings fixed effects estimation framework. Before moving on to the exact methodology used for estimation, we first discuss the following conceptual framework to our analysis. Based on Gleewe & Kremer (2006), the production function for learning outcomes can be described as -

$L = f(S, Q, C, H, I)$ - (i)

(i) herein forms the structural relationship describing learning outcomes L as a function of the various inputs into the learning process which includes – S as the number of years of schooling, Q as the school characteristics, C as the child level characteristics such as ability, H as the household characteristics such as parents attitude towards education and I as the inputs that aid the learning process such as books, tuitions. These inputs such as S, Q and I can be endogenous, as parents depending upon their attitude towards education and different child characteristics decide on the number or years of schooling, the type of school that the child attends, expenditure on books, tuitions etc. Given this S, Q and I can be further expressed as functions of the various observed and unobserved exogenous variables such as age and gender of the child, parent's education and attitude towards education, household size, highest level of education in the household, availability of basic facilities such as sanitation, electricity in the household, caste and religion of the household, household wealth etc. The reduced form version of (i) can be thus be expressed as –

$L = g(Child\ Characteristics,\ Parents\ Characteristics,\ HH\ Characteristics) - (ii)$

We estimate equation (ii) using a sibling fixed effects estimation wherein we compare children who belong to the same mother. The reason we do this is, sibling fixed effects estimation allows us to control for several factors, such as income levels, mother's female labour force participation etc, that

remain unobserved within the ASER survey and might correlate with the variable of interest. One such instance is the case of caste group, which remains unobserved in the ASER survey, households belonging to different caste groups might have different preferences or attitude towards girl child and as a result omission of the same might bias coefficient for gender.

A problem, however, that lies herein is that our dependent variable, that is performance for the individual assessment test, is an ordered variable and as a result an ordered regression might be a better specification compared to a linear one. However, the use of a non linear regression framework with fixed effects complicates interpretation and in order to keep interpretation as easy as possible, we do the following - we convert our dependent variable into a binary outcome that takes a value 1 if the child scores the highest level for the assessment test under consideration and 0 otherwise. Given that these tests are administered to children of different age groups and as a result, by the design of the tests, younger children cannot be expected to achieve the highest level. For the case of maths assessment, for instance, the highest achievement level requires completing division based tasks and division in India is usually taught to children while they are in fifth standard. Similarly, for the case of reading the highest level of the assessment tests have been designed based on what is taught in class three. To deal with this problem, we therefore focus on children who are aged 10 or above in the sample, now 10 is the age at which an average child would be in fifth standard in India. Given this, we estimate the following equation for the pooled sample of children who are aged 10 or above using sibling fixed effects in a linear framework.

$$y^j_{imt} = \beta_0 + \beta_1\, Girl_{imt} + \beta_3\, Age_{imt} + \beta_4\, Age^2_{imt} + \beta_5\, Private\ School_{imt} + \beta_6\, Tuition_{imt} + \beta_7\, Younger\ Child_{imt} + \beta\, Interactions + \delta_m + \epsilon_{imt}$$

i, herein, denotes an individual child, m denotes the mother, t denotes the time period and j denotes the assessment test. y^j_{imt} takes a value 1 if the child i achieves the highest possible level in assessment test j and 0 otherwise. Similarly, while considering performance on all the three tests together, the dependent variable y_{imt} takes a value 1 if the child achieves the highest possible level for each of the assessment tests and 0 otherwise. $Girl_{imt}$ is a dummy variable that takes a value 1 if the child is a girl child and 0 otherwise. Similarly $PrivateSchool_{imt}$ and $Tuition_{imt}$ are dummy variable that takes a value 1 if the child is enrolled into a private school and tuition respectively and 0 otherwise. δ_m represents the sibling fixed effects. $Younger\ Child_{imt}$ takes a value 1 if the child is not the oldest child and 0 otherwise. $Interactions$ denote the different interaction variables that we include in the model, these include interaction of girl dummy with the following variables - a) dummies identifying whether the individual belongs to ASER 2012 or 2014 round, b) mother's education, c) father's education, d) private school, e) tuition, f) younger child, and g) number of children. We interact girl dummy with the two time dummies in order to allow for the coefficient of the girl dummy to change over time. While the remaining interaction terms allow for a differential effect of a variable on girl's outcomes, for instance mother's level of education might have a different effect on girls outcome compared to the same for boys. Finally, we measure both mother's and father's education by dividing the same into the following four categories - a) mother's (or father's) education = 0 if years of school completed equals 0, b) mother's (or father's) education = 1 if years of school completed exceeds 0 but are less than or equal to 5, c) mother's (or father's) education = 3 if years of school completed exceeds 5 but are less than or equal to 8, and d) mother's (or father's) education = 4 if years of school completed exceeds 8.

The interest herein lies in estimating the marginal effect of being a girl on performance in assessment tests. Given the various interactions considered above and in order to ease interpretation, we define a baseline case as one wherein all the interaction terms (except for the interaction of girl dummy with the number of children [6], given that we use sibling fixed effects estimation, the num-

[6] Number of children herein refers to number of children in the age group 6 to 16, who are belong to the same mother. ASER Survey does not provide information on the exact number of children in the household and therefore a proxy has been created for the same based on those, in 5 to 16 year age group, who were assessed using the assessment tests during the survey.

ber of children at minimum therefore can take a value equal to 2) equals zero whenever $Girl_{imt}$ equals 1. Accordingly, a girl at the baseline has the following characteristics - a) mother and father both have zero years of schooling, b) is the oldest child in the household, c) does not attend private school and tuition and d) belongs to round 2009 of the survey. The marginal effect at the baseline is accordingly given by the sum of the estimated beta coefficient for the $Girl_{im}$ dummy and the interaction term $Girl * No.\ of\ Children$. Similarly, for the year 2012 and 2014 the marginal effect of being a girl at the baseline $\beta_{Girl} + \beta_{(Girl * No.\ of\ Children)} * No.\ of\ Children + \beta_{Girl*t_{12}}$ and $\beta_{Girl} + \beta_{(Girl * No.\ of\ Children)} * No.\ of\ Children + \beta_{Girl*t_{14}}$ respectively.

Descriptive Analysis

Table 2 presents estimates, from the three ASER rounds, for performance on different tests for both boys and girls in the age group 10-16 years. The numbers in the table measures the percentage of individuals at each level, for instance in the year 2009, out of all the boys in 10-16 age group, around 68.20 % could achieve the highest level (Level 4) in Reading assessment. Given this, the following observations can be made from table 2 -

a) The percentage of individuals, who are able to complete the highest level for each of the assessment tests, has fallen for both boys and girls between 2009 and 2012, while only marginal improvements can be seen for the case of english reading assessment between 2012 and 2014 for both boys and girls. A similar increase of less than 1 percentage point can also be observed for the case of reading assessment for girls between 2012 and 2014.

b) Similarly, the percentage of individuals at the lowest level for each of the tests has increased overtime across both boys and girls.

c) Gender Gap - Looking at the percentage of individuals who are in the higher most category, we see that for reading assessment in 2009 boys performed better than girls, however the same seems to have reversed over time. For boys the percentage of individuals in the top category declined to around 63% from 68% while for girls the same improved by around 1 percentage point between 2009 and 2014. The case, however, is different for maths and english reading assessment. For both maths and english reading assessment we see that the percentage of boys in the top most category exceeds the same for girls across all the three rounds and the difference between the two has increased overtime but the increase is only marginal.

Estimation Results

Descriptive analysis presented above depicts a small but widening gap between boys and girls with regard to their performance in the assessment tests conducted within the ASER survey. The question that arises next, however, is - are these gaps significant enough once we control for different individual as well household specific characteristics. In order to dwell into the same, and as already described before in the methodology section, for the pooled data from the three rounds under consideration we analyse individual's performance for these assessment tests in a siblings fixed effects framework. As before, our results in the current section are limited to children in the age group 10 and above only. Table 8 presents the results for the case wherein we consider performance on all the three assessment tests at the same time and accordingly the dependent variable takes a value 1 if the child scores the highest level for all the three assessment tests and 0 otherwise. We estimate the relationship for child's performance in table 8 for various specifications and the same are presented in different columns of table 8.

For the results presented in column 1 of table 8, following are the different explanatory variables included into the model - girl dummy, dummies for private school, tuition and younger child, age, age squared and interaction of girl dummy with the two time dummies for round 2012 and 2014. Further,

Table 2: Performance across ASER assessment tests

Year	Test	Gender	Level				
			0	1	2	3	4
2009	Reading	Male	1.43	4.90	8.58	16.89	68.20
		Female	1.67	5.41	8.23	16.81	67.88
	Maths	Male	1.50	5.28	13.15	24.01	56.06
		Female	1.73	6.35	13.99	24.74	53.19
	English	Male	6.05	10.52	12.63	24.68	46.13
		Female	6.70	10.77	13.04	24.44	45.04
2012	Reading	Male	3.10	7.98	9.63	15.51	63.78
		Female	3.45	8.41	8.88	14.05	65.22
	Maths	Male	2.19	8.44	22.27	25.14	41.97
		Female	2.49	11.26	23.03	24.90	38.31
	English	Male	7.55	10.24	17.03	24.99	40.20
		Female	9.17	10.97	17.69	24.04	38.13
2014	Reading	Male	3.99	8.35	9.77	15.06	62.83
		Female	4.00	8.31	8.68	12.94	66.07
	Maths	Male	2.59	8.34	26.19	22.70	40.18
		Female	2.74	11.04	27.31	23.42	35.49
	English	Male	7.98	8.80	18.62	24.10	40.50
		Female	9.1	9.42	19.96	22.91	38.58

we add interactions of the girl dummy with dummies for mother's level of education for specification in column 2 while for column 3 we further include interactions of girl dummy with the number of children and dummies for father's level of education, younger child, private school and tuition. Each of these specifications has been estimated using siblings fixed effects estimation. Table 9, table 10 and table 11 does similar analysis for performance on individual assessment tests of reading, maths and english reading respectively. While table 12 compares the results for the three assessment tests using our preferred specification, which is the specification used for results in column 3 from table 9, table 10 and table 11.

From column 3 in table 8 we see that the coefficient for girl dummy is statistically insignificant. Note that the marginal effect of being a girl, however, at the baseline level is given by the sum of the estimated coefficients for - a) girl dummy and, b) interaction term of girl dummy with the number of children. The estimated coefficient for the interaction term of girl dummy with the number of children is negative and statistically significant at 1% level of significance. Thus, the overall marginal effect of being a girl at the baseline in 2009 is negative, which suggests that in the year 2009 a girl at the baseline level had a lower probability of scoring the highest possible level in all the three assessment tests considered together. Further, the coefficients for the interaction terms between the girl dummy and round dummies (t_{12} & t_{14}) indicates that the disadvantage associated with girls has further increased over time. The marginal effect of being a girl at baseline for 2012 and 2014 equals the marginal effect for 2009 plus the estimated coefficient for the interaction of girl dummy with time dummy for 2012 and 2014 respectively. The coefficient for both the interaction terms are negative and statistically significant at 1% level of significance, which indicates that the gap between boys and girls at the baseline has further increased overtime.

Other than the number of children, we also interacted the girl dummy with mother's and father's level of education, dummy for whether the individual is a younger born child, dummy for private school enrolment and tuition access. We hypothesize that these factors could have a different effect on the outcome under consideration depending upon whether the individual is a boy or a girl. Educated mother's, for instance, might be encouraged to invest more into girl's education in comparison to the rest. From column 3 in table 8 we see that greater levels of mother's education is associated with a positive and significant impact on girl's overall performance. While the same is not true for father's

level of education, what we find is that an increase in father's level of education beyond upper primary level is associated with a negative impact on girl's overall performance. The same can be seen from the coefficient for the interaction between the girl dummy and father education level equal to three (defined herein as upper primary level of education or above). A reason for the same can be inferred from our further results presented in table 15 wherein we see that an increase in father's level of education is associated with a greater preference for boys towards division of household resources. From table 15 we find that increase in father's level of education leads to a fall in the probability of girl's enrolment to private school and tuition, relative to the same for boys and as discussed below both private school enrolment and tuition access are associated with a positive effect on learning.

Our above results, thus, suggests of the importance of mother's level of education in lowering the gender gap in learning and goes in accordance to our hypothesis above, the same can also be seen by comparing the marginal effect of being a girl at different levels of mother's education. To start with, at the baseline level with the number of children being equal to 2, for round 2009 the marginal effect of being a girl equalled -0.0254. This disadvantage, however, clears off with an increase in mother's level of education. For mother's education level equal to 1 the marginal effect of being a girl equals 0.0031 and it further increases to 0.0481 and 0.1082 for mother's level of education equal to 2 and 3 respectively. With regard to the distribution of mother's level of education, from table 6 we see that close to around 50% of children belong to mother's with 0 years of schooling and given the above results, these are the households wherein we see boys are performing better in comparison to girls at the baseline while the opposite is true for the rest (households with mother's level of education greater than 0) wherein we find girls performing better than boys.

With regard to other interaction terms, for a younger born child we see that the coefficient for the interaction term is positive which suggests that later born girls are at an advantage as compared to the older ones. This result is similar to that by White, Ruther & Kahn (2016), which finds that older girls with younger siblings are at a disadvantage. Finally, for the case of private school enrolment and tuition access we see that both are associated with a positive and statistically significant impact on the probability of scoring the highest level in all the three tests[7], the impact however varies in accordance to gender, as suggested by the estimated coefficient for the interaction terms between girl dummy and dummy for private school enrolment and tuition access, both of which are negative and statistically significant. Thus suggesting that both boys and girls tend to gain from private school enrolment and tuition access but the gain in learning as a result is larger for boys as compared to same for girls. In the sections ahead, using data on tuition expenditure for each child, we will show as to why the differential impact of tuition can arise between boys and girls.

Finally, to summarize the results presented above, we find that -

a) At the baseline level, girls in rural India tend to lag behind boys with regard to their performance in the assessment tests and the same disadvantage has widened over time

b) Girls with educated mother tends to perform better in comparison to boys

c) Tuition access, private school enrolment and increase in the number of children in the household tends to widen the same gap

Results for Individual Assessment Tests

For the individual assessment tests as well we observe similar results as above, although some differences can be observed for the case of reading assessment. Table 12 compares the results for the three assessment tests under consideration, column 1 of the same provides results for the case of read-

[7] The gain, from private school enrolment and tuition access, towards learning have been well established in the Indian context. See, for instance, French and Kingdon (2010), Dongre and Tewary (2015)

ing assessment test, while column 2 and 3 presents the same for maths and english reading assessment test respectively. Across all the three columns in table 12, the dependent variable takes a value 1 if the child scores the highest level for the particular assessment test under consideration and 0 otherwise. For the case of reading, we see that the estimated coefficient for the girl dummy is positive but statistically insignificant while the same for the interaction term between the girl dummy and number of children is negative and statistically significant, which implies that the overall marginal effect at the baseline is negative. Further with regard to changes overtime, we see that the estimated coefficient for the interaction between girl dummy and round 2012 and 2014 dummies are both positive and statistically significant, this suggests that girls performance in reading improved overtime. Despite this improvement the marginal effect for girls at baseline in 2012 and 2014 is still negative, which suggests that girls at the baseline continue to lack behind in reading. Further results for the case of reading are similar to the case wherein we consider all the three tests at the same time, except for the differential effect of private school enrolment for girls. We see that the estimated coefficient for the interaction of girl dummy with private school enrolment is statistically insignificant, which suggests that private school enrolment does not have any differential effect on girls for the case of reading. Other than reading, the results for both maths and english reading assessment are similar to the results for performance on the three tests considered together. Both for maths as well as english reading, girls at the baseline tend to lack in comparison to boys and the same gap seems to have widened overtime.

Differential effect of tuition across gender

From table 8 we see that the marginal effect of attending private school and tuition on the learning score varies across the two genders, with girls tending to gain less from attending a private school and tuition compared to the same for boys. What explains the same effect ? Note that in our specification private school and tuition are dummy variables that takes a value 1 if the child is enrolled into the same and 0 otherwise. Now, a reason for the above differential effect between the two gender could be that girls are being sent to a lower quality private school or tuition. The quality of both the private school and tuition remains unobserved from the ASER survey, however, one thing that can be observed for the 2014 round of survey is the actual monthly expenditure for each of the individual child enrolled into tuition. Given the lack of similar data for the case of private schools, therefore we limit our analysis to the case of tuitions only. Assuming that a better quality tuition is associated with a higher expenditure on same, we estimate the marginal effect of being a girl on the monthly amount spent on tuition. Our sample herein consists of individuals from the 2014 round, who are enrolled into tuitions and for whom the expenditure on tuitions have been reported for. Using a siblings fixed effects estimation we estimate the following equation -

$$y_{im} = \beta_0 + \beta_1 \, Girl_{im} + \beta_2 \, Age_{im} + \beta_3 \, Age_{im}^2 + \beta_4 \, Private \; School_{im} + +\beta_5 \, Younger \; Child_{im} + \delta_m + \epsilon_{im}$$

i herein identifies an individual while m identifies the mother, y_{im} measures the amount spent on tuition for individual i. The independent variables have the same interpretation as before. Table 13 presents the estimation results for the above equation for the entire sample of children of 5-16 years of age, while table 14 presents the same for the case of children who are of age 10 or above. We estimate the same relationship using different specifications. For both table 13 and 14, in column 1 we regress the amount spent on tuition only on the girl dummy, while in column 2 we include additional covariates as well. Based on our preferred specification in column 2 from both table 13 and 14 we see that the estimated coefficient for girl is dummy is both negative and statistically significant, thus suggesting that on average the expenditure incurred on tuitions for girls is less as compared to the same for the case of boys. Lesser expenditure therefore might lead to a difference in the quality of tuitions and hence the differential impact of the same for boys and girls. This thus shows that even though boys and girls might have access to similar resources but there could still be further discrimination with regard to the quality of the resources as well as spending on the same.

Difference in Division of Household Resources

In this section we analyse the differences in division of household resource between boys and girls and the changes in the same overtime. A better way to the examine the same would be to use detailed data on household division of resources across gender, however, in the absence of the same we herein focus on the difference in enrolment to private schools and tuitions in order proxy for the same effect. Given that access to either private schools or tuitions requires payment of fees, a greater enrolment of boys in the same in comparison to girls therefore represents to some extent a greater division of household resources in favour of boys.

Table 3 below presents the percentage of boys and girls, in 10-16 age group, enrolled in private school for different rounds of ASER data. Across all the three rounds we see that for boys the percentage enrolled in private schools is higher in comparison to the same for girls. We see that overtime for both boys and girls there has been an increase in the proportion of those enrolled in private schools. What is more important, however, is that overtime the gap between the two has started to widen in favour of boys. The increase in the percentage of boys enrolled into private schools has been much larger overtime as compared to the same for girls and as a result the gap of around 2.49 percentage points in 2009 has increased to 6.61 in 2014. Table 4 presents similar results for the case of tuition, wherein we see that the gap between boys and girls has widened from 2.95 percentage points in 2009 to 5.40 percentage points in 2014.

Table 3: Enrolment in Private School Across Gender

Gender	2009	2012	2014
Boys	26.25	33.54	35.85
Girls	23.76	27.89	29.24

Table 4: Enrolment in Tuitions Across Gender

Gender	2009	2012	2014
Boys	28.84	27.32	29.38
Girls	25.89	23.01	24.98

Two questions follow from the above analysis - a) are these differences significant once we control for different individual as well as household level characteristics and b) if these differences are indeed significant, then how has same evolved overtime. The estimated coefficient for the interaction between girl dummy and time dummies would capture as to whether the division of household resources have been in favour of boys or girls overtime. To examine the same we separately regress enrolment to private school and tuition on different explanatory variables, which includes child gender, age, parents level of education. Besides, we also control for sibling fixed effects in the same as well.

Table 15 presents the results from the above two estimations. In column 1 of table 15 the dependent variable is enrolment in private school and it takes a value 1 if the child is enrolled in a private school and 0 otherwise while in column 2 for the same the dependent variable is enrolment in tuition and similarly it takes a value 1 if the child is enrolled in tuition and 0 otherwise. Across all the three rounds consideration and across both private school and tuition enrolment, the estimated coefficient for girl dummy is both negative and statistically significant at 1 % level of significance. Thus suggesting that the gap between boys and girls with regard to enrolment in private schools and tuitions is indeed significant even after controlling for various individual as well as household level characteristics and that the household division of resources seems to favour boys than girls. Also, we see that the absolute value of the same has increased over time, that is, the probability of a girl being enrolled in a private school or a tuition has further reduced overtime. This suggests that the possibility of a girl being discriminated against with regard to household division of resources has further increased over

time.

Role of Culture ?

In this section we analyse the role of culture and attitudes towards women in explaining gender gaps in learning. The role culture in explaining these gaps has been documented in past as well (see Guiso (2008), Pope & Syndor (2008)), this section therefore explores as to whether similar results hold for Indian states or not. In order to do so we use the Female Empowerment Index[8] for Indian states by Mckinsey Global (MGI 2015), table 5 describes the performance for different states in India on the index. Based on the same we divide our sample into two - first comprising of states which score less than (or equal to) the average for India and second comprising of those that score more than the average. The first group therefore consists of states wherein the society is more biased towards males as compared to states in the second group. The first group consists of the following states - Bihar, Jharkhand, Assam, Uttar Pradesh, Madhya Pradesh, Arunachal Pradesh, Tripura, Orissa, Rajasthan, Nagaland, Haryana and West Bengal; while the second group consists of the remaining states in India. For both the groups we separately estimate children's performance on all the three tests combined using a siblings fixed effects estimation. And as before our dependent variable takes a value 1 if the child scores the highest level in all the three assessment tests and 0 otherwise and our sample is restricted to children who are in the age group 10 or above. The results from the estimation for both the groups are presented in table 16, column 1 presents the results for the first division while column 2 presents the same for the second division. Comparing across columns in table 16 we see that the marginal effect of being a girl is larger in column 1 compared to the same in column 2, for instance at the baseline level the marginal effect of being a girl equalled -0.0486 for the case of states belonging to the first group while the same for the states in the second group equalled -0.0314. Besides for the first group we see that there has been a rise in the gender gap overtime however no such effect could be seen for the case of the second group.

Conclusion

The above analysis provides evidence for gender gap in learning from rural India. What we find is the following - a) for children belonging to mother's with zero years of education, boys tend to perform better for all the three tests for learning conducted under the ASER survey, b) an increase in mother's years of education is associated with girls performing better as compared to boys. We also find that the gap between boys and girls has widened overtime, in the post RTE period, in favour of boys for the case of maths and english reading assessment while for the case of reading girls have gained overtime as compared to boys. In terms of what explains this differential trend, a possible reason could be the *No Detention Policy* under the RTE Act. As per the same children cannot be detained or held back to repeat a class till they reach 14 years of age. What this means is that learning isn't a criteria any more for progression to the next grade as compared to a traditional setup wherein a child's progression to the next level is dependent upon his or her level of learning. An implication of the same could be that parents can now discriminate between children with regard to the skills they acquire, for instance boys could be made to spend more time in doing maths and english, which parents might perceive as being more useful in the future labour market. Additionally, our study also points towards the differences in division of household resources between boys and girls, we demonstrate the same using tuition expenditure and the probability of being enrolled in a tuition or a private school. What we find, using our analysis of enrolment to tuitions or a private school, is that not only do boys tend to gain a larger share of the household resources but the same has moved in favour of boys overtime.

For the question that we posed in the title - Dakar, are we there yet? That is, have India been

[8] The Female Empowerment Index (MGI (2005)) is based on the following ten dimensions - a) Labour Force Participation, b) Participation in Professional and Technical Jobs, c) Participation in Leadership Positions, d) Unpaid Work, e) Family Planning, f) Maternal Mortality, g) Enrolment Level, h) Sex ratio at birth, i) Child Marriage, and j) Domestic Violence. The index takes a value between 0 and 1 and higher values for the index represents gender parity.

able to close in on the gender gaps in learning or not. Our analysis suggests that in the post RTE period, except for the case of reading, the gap between boys and girls at the baseline, with regard to both maths and english reading assessment as well as distribution of household resources, have widened overtime in favour of boys. This implies that a move further away from the said objectives of gender equality in educational outcomes by 2015 under the Dakar Framework. Further, what is more terrible is the fact that the same has been happening despite having major educational reforms such as the Sarva Sikhsa Abhiyan and the RTE Act that were implemented towards bridging gaps in educational outcomes. Although progress can be seen in terms of a reduction of gender gaps in enrolment on account of these reforms but for learning greater efforts are needed. Given this, what is needed is a reorientation of the existing educational strategy for India towards ensuring a greater equality in terms of learning across gender.

References

ASER (2014). Annual Status of Education Report : Trends Over Time. ASER, New Delhi.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. Journal of Econometrics, 121(1), 343-375.

Azam, M., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. Economic Development and Cultural Change, 61(2), 335-367.

Bedard, K., & Cho, I. (2010). Early gender test score gaps across OECD countries. Economics of Education Review, 29(3), 348-363.

Berman, E., Lang, K., & Siniver, E. (2003). Language-skill complementarity: returns to immigrant language acquisition. Labour Economics, 10(3), 265-290.

Bleakley, H., & Chin, A. (2008). What holds back the second generation? The intergenerational transmission of language human capital among immigrants. Journal of Human Resources, 43(2), 267-298.

Census (2001). Census of India, Government of India.

Dongre, A., and Vibhu, T.. "Impact of Private Tutoring on Learning Levels." Economic & Political Weekly 50.41 (2015): 73.

Doris, A., O'Neill, D., & Sweetman, O. (2013). Gender, single-sex schooling and maths achievement. Economics of Education Review, 35, 104-119.

French, R. and Kingdon, G. The relative effectiveness of private and government schools in Rural India: Evidence from ASER data. No. 10-03. Department of Quantitative Social Science-UCL Institute of Education, University College London, 2010.

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. American Economic Journal: Applied Economics, 2(2), 210-240.

Glewwe, P., & Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. Handbook of the Economics of Education, 2, 945-1017.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. Science, 320(5880), 1164.

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. Journal of economic literature, 46(3), 607-668.

Joensen, J. S., & Nielsen, H. S. (2009). Is there a causal effect of high school math on labor market Outcomes?. Journal of Human Resources, 44(1), 171-198.

Klasen, S. (2002). Low schooling for girls, slower growth for all? Cross country evidence on the effect of gender inequality in education on economic development. The World Bank Economic Review, 16(3), 345-373.

Klasen, S., & Lamanna, F. (2008). The Impact of Gender Inequality in Education and Employment on Economic Growth in Developing Countries: Updates and Extensions (No. 175). Ibero-America Institute for Economic Research.

MHRD (2014). Educational Statistics at a glance. Ministry of Human Resource and Development, Government of India.

MGI (2015). The Power of Parity, Advancing Women's Parity in India. Mckinsey Global Institute.

Nollenberger, N., Rodríguez-Planas, N., & Sevilla, A. (2016). The Math Gender Gap: The Role of Culture. The American Economic Review, 106(5), 257-261.

Pope, D. G., & Sydnor, J. R. (2010). Geographic variation in the gender differences in test scores. The Journal of Economic Perspectives, 24(2), 95-108.

Rose, H., & Betts, J. R. (2004). The effect of high school courses on earnings. Review of Economics and Statistics, 86(2), 497-513.

Schultz, T. P. (2002). Why governments should invest more to educate girls. World Development, 30(2), 207-225.

White, G., Ruther, M., & Kahn, J. (2016). Educational Inequality in India: An Analysis of Gender Differences in Reading and Mathematics. IHDS Working Paper 2016-2.

Xin, Ma (2008). Paper commissioned for the EFA Global Monitoring Report 2008, Education for All by 2015: will we make it? . UNESCO.

Table 5: Table

Group	State	Index Score	State	Index Score
First Group	Bihar	0.42	Jharkhand	0.46
	Assam	0.47	Uttar Prad.	0.49
	Madhya Prad.	0.49	Tripura	0.51
	Arunachal Prad.	0.50	Orissa	0.51
	Rajasthan	0.52	Nagaland	0.52
	Haryana	0.53	West Bengal	0.54
Second Group	JnK	0.55	Manipur	0.55
	Chattisgarh	0.55	Gujarat	0.56
	Uttarakhand	0.57	Punjab	0.59
	Maharashtra	0.59	Andhra Prad.	0.59
	Karnataka	0.59	Puducherry	0.59
	Tamil Nadu	0.60	Himachal Prad.	0.63
	Sikkim	0.64	Goa	0.64
	Kerala	0.67	Meghalaya	0.69
	Mizoram	0.70		

Index Score represents the Female Empowerment Index (Source - MGI 2015). Higher values represents greater gender parity

Table 6: Summary Statistics Round Wise

	Round 2009				Round 2012				Round 2014			
Variable	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max
Girl	0.46	0.50	0	1	0.49	0.50	0	1	0.51	0.50	0	1
No. of Children	2.86	0.97	2	10	2.78	0.94	2	12	2.75	0.91	2	8
Tuition	0.27	0.45	0	1	0.25	0.43	0	1	0.27	0.44	0	1
Private School	0.25	0.43	0	1	0.31	0.46	0	1	0.32	0.47	0	1
Age	12.50	1.93	10	16	12.48	1.89	10	16	12.46	1.87	10	16
Mother Ed=0	0.53	0.50	0	1	0.52	0.50	0	1	0.50	0.50	0	1
Mother Ed=1	0.18	0.38	0	1	0.17	0.37	0	1	0.16	0.37	0	1
Mother Ed=2	0.14	0.34	0	1	0.14	0.35	0	1	0.15	0.36	0	1
Mother Ed=3	0.16	0.36	0	1	0.17	0.38	0	1	0.19	0.39	0	1
Father Ed=0	0.30	0.46	0	1	0.27	0.44	0	1	0.26	0.44	0	1
Father Ed=1	0.17	0.37	0	1	0.17	0.37	0	1	0.16	0.36	0	1
Father Ed=2	0.17	0.38	0	1	0.18	0.38	0	1	0.17	0.38	0	1
Father Ed=3	0.36	0.48	0	1	0.39	0.49	0	1	0.41	0.49	0	1
N	208729				166103				154246			

Table 7: Summary Statistics (Combined)

Variable	Mean	Std Dev	Min	Max
Girl	0.48	0.50	0	1
No of Sibling	2.80	0.94	2	12
Tuition	0.27	0.44	0	1
Private School	0.29	0.45	0	1
Age	12.48	1.90	10	16
Mother Ed=0	0.52	0.50	0	1
Mother Ed=1	0.17	0.37	0	1
Mother Ed=2	0.14	0.35	0	1
Mother Ed=3	0.17	0.38	0	1
Father Ed=0	0.28	0.45	0	1
Father Ed=1	0.16	0.37	0	1
Father Ed=2	0.17	0.38	0	1
Father Ed=3	0.38	0.49	0	1
N	529078			
(ASER 2014)				
Tuition Amount	234.02	266.48	1	5000
N	34114			

Table 8: Results for performance on all three assessment tests

	(1) Highest Level	(2) Highest Level	(3) Highest Level
Girl	-0.0144***	-0.0317***	0.0009
	(0.0028)	(0.0033)	(0.0076)
Girl * t_{12}	-0.0128**	-0.0130**	-0.0138**
	(0.0043)	(0.0044)	(0.0044)
Girl * t_{14}	-0.0163***	-0.0178***	-0.0181***
	(0.0046)	(0.0046)	(0.0046)
Private School	0.0413***	0.0410***	0.0461***
	(0.0035)	(0.0036)	(0.0040)
Tuition	0.103***	0.103***	0.111***
	(0.0040)	(0.0041)	(0.0045)
Younger Child	-0.0474***	-0.0473***	-0.0611***
	(0.0026)	(0.0026)	(0.0032)
Girl * Mother Education = 1		0.0275***	0.0285***
		(0.0052)	(0.00533)
Girl * Mother Education = 2		0.0364***	0.0418***
		(0.0058)	(0.0061)
Girl * Mother Education = 3		0.0471***	0.0571***
		(0.0056)	(0.0063)
Girl * Father Education = 1			0.0061
			(0.0057)
Girl * Father Education = 2			0.0019
			(0.0057)
Girl * Father Education = 3			-0.0167**
			(0.0051)
Girl * Younger Child			0.0283***
			(0.0037)
Girl * Private School			-0.0115**
			(0.0042)
Girl * Tuition			-0.0221***
			(0.0047)
Girl * No. of Children			-0.0127***
			(0.0019)
Observations	529078	529078	529078
R^2	0.190	0.190	0.191

Standard errors clustered at sibling level in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable takes a value 1 if the child scores the highest level in all the three assessment tests and 0 otherwise. Each specification has been estimated with siblings fixed effects. Additional controls include child age, age squared and a constant.

Table 9: Results for performance on reading tests

	(1) Highest Level in Reading Test	(2) Highest Level in Reading Test	(3) Highest Level in Reading Test
Girl	-0.0092**	-0.0241***	0.0074
	(0.0028)	(0.0033)	(0.0080)
Girl * t_{12}	0.0150***	0.0149***	0.0130**
	(0.0044)	(0.00443)	(0.00443)
Girl * t_{14}	0.0161***	0.0149**	0.0130**
	(0.0046)	(0.0046)	(0.0046)
Private School	0.0460***	0.0457***	0.0469***
	(0.0034)	(0.0034)	(0.0039)
Tuition	0.0719***	0.0714***	0.0774***
	(0.0039)	(0.0039)	(0.0044)
Younger Child	-0.0312***	-0.0311***	-0.0449***
	(0.00274)	(0.00274)	(0.00331)
Girl * Mother Education = 1		0.0263***	0.0225***
		(0.00533)	(0.00549)
Girl * Mother Education = 2		0.0336***	0.0302***
		(0.00554)	(0.00591)
Girl * Mother Education = 3		0.0349***	0.0319***
		(0.00500)	(0.00572)
Girl * Father Education = 1			0.0139*
			(0.00626)
Girl * Father Education = 2			0.0130*
			(0.00606)
Girl * Father Education=3			0.0023
			(0.00535)
Girl * Younger Child			0.0286***
			(0.00373)
Girl * Private School			-0.00245
			(0.00412)
Girl * Tuition			-0.0166***
			(0.00462)
Girl * No of Children			-0.0156***
			(0.00200)
Observations	529078	529078	529078
R^2	0.202	0.202	0.203

Standard errors clustered at sibling level in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable takes a value 1 if the child scores the highest level in reading tests and 0 otherwise. Each specification has been estimated with siblings fixed effects. Additional controls include child age, age squared and a constant.

Table 10: Results for performance on maths tests

	(1) Highest Level in Maths Test	(2) Highest Level in Maths Test	(3) Highest Level in Maths Test
Girl	-0.0214***	-0.0396***	-0.0146
	(0.0029)	(0.0034)	(0.0081)
Girl * t_{12}	-0.0185***	-0.0188***	-0.0198***
	(0.0046)	(0.0046)	(0.0046)
Girl * t_{14}	-0.0375***	-0.0391***	-0.0398***
	(0.0049)	(0.0048)	(0.0049)
Private School	0.0423***	0.0420***	0.0474***
	(0.0036)	(0.0036)	(0.0041)
Tuition	0.107***	0.106***	0.114***
	(0.0042)	(0.0042)	(0.0047)
Younger Child	-0.0442***	-0.0441***	-0.0608***
	(0.0028)	(0.0028)	(0.0034)
Girl * Mother Education = 1		0.0273***	0.0263***
		(0.0055)	(0.0057)
Girl * Mother Education = 2		0.0383***	0.0395***
		(0.0060)	(0.0064)
Girl * Mother Education = 3		0.0512***	0.0557***
		(0.0056)	(0.0064)
Girl * Father Education = 1			0.0044
			(0.0062)
Girl * Father Education = 2			0.0117
			(0.0062)
Girl * Father Education = 3			-0.0052
			(0.0055)
Girl * Younger Child			0.0342***
			(0.0039)
Girl * Private School			-0.0118**
			(0.0044)
Girl * Tuition			-0.0215***
			(0.0049)
Girl * No. of Children			-0.0126***
			(0.0020)
Observations	529078	529078	529078
R^2	0.175	0.176	0.177

Standard errors clustered at sibling level in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable takes a value 1 if the child scores the highest level in maths tests and 0 otherwise. Each specification has been estimated with siblings fixed effects. Additional controls include child age, age squared and a constant.

Table 11: Results for performance on english reading tests

	(1) Highest Level in English Test	(2) Highest Level in English Test	(3) Highest Level in English Test
Girl	-0.0127***	-0.0297***	0.0104
	(0.0028)	(0.0033)	(0.0079)
Girl * t_{12}	-0.0139**	-0.0141**	-0.0147***
	(0.0044)	(0.0044)	(0.0044)
Girl * t_{14}	-0.0188***	-0.0202***	-0.0202***
	(0.0047)	(0.0047)	(0.0047)
Private School	0.0743***	0.0740***	0.0812***
	(0.0035)	(0.0035)	(0.0040)
Tuition	0.104***	0.104***	0.110***
	(0.0040)	(0.0040)	(0.0045)
Younger Child	-0.0433***	-0.0432***	-0.0551***
	(0.0026)	(0.0026)	(0.0032)
Girl * Mother Education = 1		0.0269***	0.0280***
		(0.0053)	(0.0055)
Girl * Mother Education = 2		0.0376***	0.0436***
		(0.0058)	(0.0062)
Girl * Mother Education = 3		0.0446***	0.0563***
		(0.0052)	(0.0061)
Girl * Father Education = 1			0.0046
			(0.0059)
Girl * Father Education = 2			0.0027
			(0.0059)
Girl * Father Education = 3			-0.0203***
			(0.0052)
Girl * Younger Child			0.0244***
			(0.0037)
Girl * Private School			-0.0165***
			(0.0042)
Girl * Tuition			-0.0189***
			(0.0047)
Girl * No. of Children			-0.0138***
			(0.0019)
Observations	529078	529078	529078
R^2	0.208	0.209	0.210

Standard errors clustered at sibling level in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable takes a value 1 if the child scores the highest level in english reading tests and 0 otherwise. Each specification has been estimated with siblings fixed effects. Additional controls include child age, age squared and a constant.

Table 12: Comparison across tests

	(1) Highest Level in Reading Test	(2) Highest Level in Maths Test	(3) Highest Level in English Test
Girl	0.00741	-0.0146	0.0104
	(0.0080)	(0.0081)	(0.0079)
Girl * t_{12}	0.0130**	-0.0198***	-0.0147***
	(0.0044)	(0.0046)	(0.0044)
Girl * t_{14}	0.0130**	-0.0398***	-0.0202***
	(0.0046)	(0.0049)	(0.0047)
Girl * Mother Education = 1	0.0225***	0.0263***	0.0280***
	(0.0054)	(0.0057)	(0.0055)
Girl * Mother Education = 2	0.0302***	0.0395***	0.0436***
	(0.0059)	(0.0064)	(0.0062)
Girl * Mother Education = 3	0.0319***	0.0557***	0.0563***
	(0.0057)	(0.0064)	(0.0061)
Girl * Father Education = 1	0.0139*	0.0044	0.0046
	(0.0062)	(0.0062)	(0.0059)
Girl * Father Education = 2	0.0130*	0.0117	0.0027
	(0.0060)	(0.0062)	(0.0059)
Girl * Father Education = 3	0.00231	-0.00529	-0.0203***
	(0.0053)	(0.0055)	(0.0052)
Girl * Younger Child	0.0286***	0.0342***	0.0244***
	(0.0037)	(0.0039)	(0.0037)
Girl * Private School	-0.00245	-0.0118**	-0.0165***
	(0.0041)	(0.0044)	(0.0042)
Girl * Tuition	-0.0166***	-0.0215***	-0.0189***
	(0.0046)	(0.0049)	(0.0047)
Girl * No. of Children	-0.0156***	-0.0126***	-0.0138***
	(0.0020)	(0.0020)	(0.0019)
Private School	0.0469***	0.0474***	0.0812***
	(0.0039)	(0.0041)	(0.0040)
Tuition	0.0774***	0.114***	0.110***
	(0.0044)	(0.0047)	(0.0045)
Younger Child	-0.0449***	-0.0608***	-0.0551***
	(0.0033)	(0.0034)	(0.0032)
Observations	529078	529078	529078
R^2	0.203	0.177	0.210

Standard errors clustered at sibling level in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable takes a value 1 if the child scores the highest level in the tests and 0 otherwise, first column considers test for reading, second column considers test for maths and the third column considers test for english reading. Each specification has been estimated with siblings fixed effects. Additional controls include child age, age squared and a constant.

Table 13: Differences in Tuition Expenditure Across Gender

	(1) Tuition Expenditure	(2) Tuition Expenditure
Girl	-10.53***	-19.83***
	(2.362)	(2.295)
Age		-34.31***
		(3.092)
Age Square		2.288***
		(0.149)
Younger Child		-13.63***
		(2.648)
Private School		30.10***
		(8.781)
Constant	212.5***	300.8***
	(1.102)	(16.60)
Observations	52843	52843
R^2	0.001	0.098

Standard errors clustered at sibling level in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable measures the tuition expenditure for each child. The sample consists of children who are enrolled into tuitions. Each specification has been estimated with siblings fixed effects.

Table 14: Differences in Tuition Expenditure Across Gender (10 and above)

	(1) Tuition Expenditure	(2) Tuition Expenditure
Girl	-14.21***	-25.01***
	(3.595)	(3.321)
Child Age		-141.7***
		(13.72)
Age Square		6.517***
		(0.553)
Younger Child		-10.29*
		(4.677)
Private School		39.92**
		(12.88)
Constant	240.6***	969.0***
	(1.715)	(85.94)
Observations	34014	34014
R^2	0.001	0.109

Standard errors clustered at sibling level in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable measures the tuition expenditure for each child. The sample consists of children who are enrolled into tuitions and are aged 10 or above. Each specification has been estimated with siblings fixed effects.

Table 15: Enrolment in Private Schools and Tuition

	(1) Private School	(2) Tuition
Girl	-0.0075	-0.0095
	(0.0062)	(0.0055)
Girl * t_{12}	-0.0165***	-0.0143***
	(0.0036)	(0.0031)
Girl * t_{14}	-0.0285***	-0.0216***
	(0.0038)	(0.0034)
Girl * Mother Education = 1	0.0184***	0.0111**
	(0.0043)	(0.0037)
Girl * Mother Education = 2	0.0177***	0.0155***
	(0.0050)	(0.0042)
Girl * Mother Education = 3	0.0206***	0.0237***
	(0.0051)	(0.0045)
Girl * Father Education = 1	-0.0079	0.0021
	(0.0046)	(0.0040)
Girl * Father Education = 2	-0.0185***	-0.0067
	(0.0047)	(0.0040)
Girl * Father Education = 3	-0.0281***	-0.0203***
	(0.0043)	(0.0037)
Girl * Younger Child	-0.0069*	0.0182***
	(0.0028)	(0.0026)
Girl * No. of Children	-0.0051**	-0.0099***
	(0.0016)	(0.0014)
Younger Child	-0.0095***	-0.0145***
	(0.0025)	(0.0022)
Observations	529078	529078
R^2	0.016	0.030

Standard errors clustered at sibling level in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable takes a value 1 if the child scores the highest level in reading tests and 0 otherwise. Each specification has been estimated with siblings fixed effects. Additional controls include child age, age squared and a constant.

Table 16: Heterogeneity Across States

	(1) Highest Level in all 3 tests	(2) Highest Level in all 3 tests
Girl	-0.0398***	0.0279
	(0.0090)	(0.0143)
Girl * t_{12}	-0.0179***	-0.00157
	(0.0053)	(0.0074)
Girl * t_{14}	-0.0193***	-0.00398
	(0.0057)	(0.0077)
Girl * Mother Education = 1	0.0220**	0.0223*
	(0.0067)	(0.0089)
Girl * Mother Education = 2	0.0361***	0.0225*
	(0.0082)	(0.0095)
Girl * Mother Education = 3	0.0443***	0.0281**
	(0.0088)	(0.0097)
Girl * Father Education = 1	0.0051	0.0074
	(0.0068)	(0.0098)
Girl * Father Education = 2	-0.0015	0.0133
	(0.0068)	(0.0102)
Girl * Father Education = 3	-0.0242***	0.00683
	(0.0061)	(0.0092)
Girl * Younger Child	0.0431***	0.00813
	(0.0045)	(0.00643)
Girl * Private School	-0.0224***	0.00285
	(0.0052)	(0.0071)
Girl * Tuition	-0.0130*	-0.00426
	(0.0055)	(0.0090)
Girl * No. of Children	-0.0044*	-0.0157***
	(0.0022)	(0.0040)
Private School	0.0428***	0.0376***
	(0.0049)	(0.0068)
Tuition	0.125***	0.0639***
	(0.0053)	(0.0086)
Younger Child	-0.0571***	-0.0599***
	(0.0039)	(0.0056)
Observations	322831	202702
R^2	0.187	0.206

Standard errors clustered at sibling level in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The dependent variable takes a value 1 if the child scores the highest level in all the three tests and 0 otherwise. The first column presents the results for group 1 states while column 2 presents the same for group 2. Each specification has been estimated with siblings fixed effects. Additional controls include child age, age squared and a constant.