# Peer Gender Effects: Evidence from India


Vanshika Chaudhary

PhD Candidate, Economics

Indian Institute of Management Indore

Ajay Sharma

Associate Professor, Economics

Indian Institute of Management Indore

Corresponding Author: Vanshika Chaudhary, Indian Institute of Management Indore, Prabandh Shikhar, Rau-Pithampur Road, Indore (Madhya Pradesh)-453556; email: f20vanshikac@iimidr.ac.in; Ph: +91-9584292070

# Peer Gender Effects: Evidence from India


## Abstract

In this paper, we investigate the effect of peer gender composition for students aged 9-10 years in India using data from Young Lives study, on different cognitive abilities of girls and boys. We exploit the de facto random assignment of students across classes, conditional on lagged ability and school fixed effects. We find that an increase in the proportion of female students in the classroom causes females' math abilities to improve, while it causes boys' English abilities to improve. When we exclude single-sex schools from the sample, we find that girls also experience a positive peer gender effect on English abilities. We find that our results are driven by public-school going students, who study in schools that do not offer instruction in English medium and schools that do not provide basic amenities like gender-segregated toilets for girls. Finally, we consider various channels through which peer gender effects cause improvement and heterogeneity in improvement in cognitive abilities, by gender.



*Keywords: peer gender effect, cognitive abilities, gender gap in mathematics, Young Lives, single-sex schools, primary school education India*

*JEL Codes: I21, I25, J16*

## 1. Introduction

The effect of peers has long been established as a powerful determinant of why we are the way we are. Academic enquiry has revealed that peer influences play a major role in generating a vast set of individuals' outcomes, including cognitive and noncognitive abilities, test scores, college education, choice of career, use of drugs, teenage pregnancy, gender attitudes, criminality and delinquency, among others (Sacerdote, 2011). The economics literature on the peer effects in education, in particular, has been traced back to the Coleman report (1968) (Epple & Romano, 2011; Moffitt, 2001; Sacerdote, 2011), and has recently been receiving ever increasing attention (Sacerdote, 2011). Yet, most of these studies are predominantly and disproportionately based on data from developed countries (Borbely et al., 2022; Hu, 2015), even as they explore various aspects of peer composition effects including that of peers' outcomes, peer current behaviours, and peers' backgrounds on an individual's educational outcomes. In this paper, we examine what happens to the cognitive abilities of girls and boys when we vary the gender composition of their peers in the classroom in a developing country. This question takes an important position from the perspective of informing education policy in a developing country context given that it allows us to investigate externalities of encouraging and increasing the participation of girls in the formal education system for both girls and boys, and leverage the same.

This paper studies the impact of gender composition of classroom peers of students at the primary school level, on their cognitive abilities using the Young Lives (YL) study data in India. Two aspects of this dataset make it novel. One, it is one of the few longitudinal datasets that track Indian children, collating extensive background information on these children, including demographic characteristics and tracking ability changes using child, household and community level surveys. Two, this survey also includes a special round of School Survey which provides comprehensive data on various aspects of the school life of the YL child which can be directly mapped to other data on the YL child available from the regular rounds of the survey. This allows us to exploit a unique and rich dataset on the YL children for our study, from the primary school level.

Primary school education, also called the "biggest equaliser" (Chandra, 2017), is a critical stage to study the impact of various factors, especially from the perspective of cognitive development. Studies have shown that it is a critical period in a child's development as these are the very years "when the gender gap (in mathematical abilities) first opens up" (Doris et al., 2013; Fryer & Levitt, 2010; LoGerfo et al., 2006; Robinson & Lubienski, 2011). However, the primary school education delivery system in India is anything but adequate, even a decade and a half after the Right to Education Act of 2009 made education a fundamental right in India, promising free and compulsory schooling for all children aged between six and fourteen years. The ASER (2023) highlights how while enrolment numbers at the primary school level have seemingly gone up, attendance numbers have certainly not improved, and similarly, foundational skills at literacy and numeracy have been dipping (Gazta & Jadhav, 2023). Further, high drop-out rates, poor teacher quality, increase in multigrade classrooms, and regressive caste and gender norms plague the institution of primary education in India (Gazta & Jadhav, 2023). Given the criticality of this stage in the cognitive development of the child (Knudsen et al., 2006; Kautz et al., 2014), and the urgent need to improve the primary school education delivery system in India, as Borbely et al. (2022) note, there is a huge potential for leveraging the role of peer composition on students' outcomes at the primary school level in education policy. This might be especially important for girls as the gender gap in abilities, especially mathematical and quantitative abilities, which opens up at this stage, only widens over time (Rakshit & Sahoo, 2024).

Peer effect estimations are not straight forward because of the endogeneity of peer group composition. Individuals may self-select into groups based on pre-determined characteristics such as gender, confounding peer gender effect with selection bias. In our analysis, accordingly, we address the issue of endogeneity and identification by using the idiosyncratic variation in the gender composition of classrooms across different schools. We argue that after controlling for school fixed effects, the variation in the gender composition of classrooms is effectively exogenous, arising on account of, one, the natural biological demographic variation across the population (Dewan et al., 2024; Hoxby, 2000), and two, because of the de facto random assignment of students across classes in a school (Eisenkopf et al., 2015). Further, our data allows us to control for the past abilities of students, and that effectively controls against selection based on individual-level quality variations amongst students as well as unobserved past inputs into the ability production function (Hanushek, 1979; Singh, 2015).

Our analyses yield a positive and robust peer gender effect from increasing the proportion of girl peers in the classroom on the cognitive abilities of the child, for both boys and girls. However, we observe a distinct pattern of gender heterogeneity of robust peer gender effects by ability. We find that corresponding to a ten percentage point increase in the proportion of girls in the class, for girls, there is, on an average, a causal improvement in their mathematical and quantitative abilities by 10.4 percent of a standard deviation, while for boys, for a similar change in peer gender composition, we find improvement in their English abilities to the tune of about 14 percent of a standard deviation and no effect on their math abilities, on an average. When we exclude the single sex schools from our sample to account for endogenous differences between single sex and coeducational schools, we further find that girls see an improvement in both Math as well as English abilities by 11.6 and 11.5 percents of a standard deviation, respectively, as a result of a ten percentage point increase in the proportion of females in classrooms in coeducational schools, while the results for males remain much the same. For both males and females, we find no peer gender composition effect on Telugu abilities. Our estimates strongly suggest that there exist distinct mechanisms through which peer gender composition positively impact boys' versus girls' outcomes.

In literature, we find support for both uniform peer gender composition effects for males and females, as well as asymmetric peer gender composition effects by gender, and studies explore various mechanisms through which these peer gender composition effects actualize.

Jackson (2012) finds a positive effect of single-sex schools only for girls, and only for those girls that have a strong preference for gender segregated education, using random assignment of girls in Trinidad and Tobago to secondary schools, while recording no effect for boys. Eisenkopf et al. (2015) also note a positive effect of single-sex schooling on the math performance of Swiss high-school female students, but no effect on performance in German. They attribute the improvement to improvement in self-perception of academic agency in girls in single-sex schools on mathematics. Riordan (1985) finds that single-sex education significantly benefits girls but not boys at all. Bryk et al. (1993) report positive academic as well as social and personal development for girls because of girls-only schools. Booth et al. (2013) through experimental evidence show that random assignment of girls to single-sex economics tutorials improves girls' outcomes as compared to those of girls assigned to mixed gender tutorials, even as girls don't tend to choose their majors differently across single-sex and mixed tutorials. They record no effect for boys. They argue that the effect might be the impact of differences in engagement in competition due to the presence of the opposite sex. Borbely et al. (2022) propose that reduction in stereotype threats from smaller number of boys in the classroom, and fewer disruptions because of females' tendencies of exhibiting fewer disruptive behaviours, lead to their results which show an improvement in math outcomes for girls in secondary schools in Ethiopia on account of an increase in the proportion of girls in the classroom, while they too record no effects for boys. Lee and Lockheed (1990) attribute the positive effects of single-sex schools on girls in the 9th grade in Nigeria to mitigation of stereotype threats from having more girls in the classroom. Lu and Anderson (2015) show that being surrounded by five more girls as opposed to five more boys improves girls' test scores but not those of boys in Chinese middle schools where students are randomly assigned to their desks in classrooms.

On the other hand, Anelli and Peri (2019) find, in line with Park et al. (2013) that single-sex settings only impact males, and they attribute it to the idea that males are more likely to both form and be influenced by larger same-sex networks while females tend to prefer smaller networks (Stehle et al., 2013). Stevenson (2015) also finds that girls are not influenced by being exposed to a higher number of high risk inmates in juvenile prisons while boys are, attributing it to the differences in norms of intra-sex social networking. Coleman (1968) found negative peer gender effect of gender mixing for boys. Later studies like Black et al. (2013), Briole (2021), Hill (2015), Jackson (2021), and Mael et al. (2005) similarly recorded negative returns for boys on account of increased exposure to girls in different educational environments. They all argue that prevalent dating norms that predominantly affect boys, lead to their being distracted by the presence of girls in the peer groups leading to negative impacts on their academic outcomes. Some studies have even recorded no or weak effects (Antecol et al., 2016; Getik & Meier, 2022; Oosterbeek & Van Ewijk, 2014).

In contrast, there are a plethora of studies that show that peer gender effects of both increased same sex exposure as well as opposite sex exposure can be positive, and for both males as well as females. Gong et al. (2019) find that both males' and females' outcomes improve across domains of cognitive ability with the increase in the proportion of females in the classroom, among the high school students in China. They argue that the effects are channelled through improved classroom environment, more positive shifts in teacher behaviour, and an increase in student efforts from social interactions. Dewan et al. (2024) also report positive effects of increased classroom female peers on both girls and boys in secondary school in India, noting that it is higher for boys than girls and attribute it to lowering of the tendency of disruptions. Hoxby (2000b) use idiosyncratic variation in gender across cohorts in elementary schools and find that peer test scores improve for both boys and girls with an increase in the number of girls in the classroom. On the other hand, Doris et al. (2013) find that single-sex schools benefit both boys and girls by improving their math abilities, and reason that these results are obtained on account of differential attitudes to competition and learning as well as reduction in stereotype salience in single-sex schools. Similarly Billger (2009), and Park et al. (2013) also find that both males and females benefit from single-sex education in their mathematical abilities.

Based on our survey of literature, we categorise all proposed mechanisms of peer gender effects under six heads – one, classroom disruption, two, social interaction, three, competitive environment, four, teacher behaviour, five, approaches to learning, and six, gender stereotyping. We consider each of these mechanisms and propose that our results for girls' benefitting at mathematical abilities are most likely driven by reduction in stereotype salience, while our findings for boys' improvement in language abilities likely originate from the dynamics of social interaction. Both boys and girls also likely benefit from an overall improvement in the classroom environment. We note that we are unable to comment on the role of teacher behaviour in driving these results.

As pointed out earlier, previous literature on peer gender effects predominantly deals with the context of developed countries. Dewan et al. (2024) eloquently illustrate this point by highlighting a number of studies in the context of the US, France, Switzerland, Israel, Italy, Norway, Austria, Sweden, Denmark, Ireland and South Korea. There are a few studies in the Chinese context including Lu and Anderson (2015), Hu (2015), Gong et al. (2019), and Luo and Yang (2023) at the middle school level, one in the Nigerian context (Lee & Lockheed, 1990) at the middle school level, one in the Ethiopian context (Borbely et al., 2022) also at the middle school level, and only one in the Indian context, Dewan et al. (2024) that studies the secondary school students with an average age of 15-16 years. To the best of our knowledge, this is one of the only studies in the context of a developing country like India, estimating the peer gender effects at the primary school level.

While Hu (2015) finds that peer gender effects mainly impact boys, across math, English and Chinese, Lu and Anderson (2015) & Borbely et al. (2022) find no effect for males and that peer gender effects mainly work for girls, across math and languages. On the other hand, Gong et al. (2019) and Dewan et al. (2024) find positive peer gender effects for both boys and girls. Thus, our study is distinct from all the other developing economies' studies on peer gender effects because one, we study primary school children, and two, we find that not only do both boys and girls benefit from an increased proportion of females in the classroom, but that they benefit in distinctly gendered ways which is evidenced by the positive effects of peer gender composition accruing asymmetrically by discipline to both the genders.

Our findings have a direct bearing on education policy because our results suggest that there exists a positive peer gender effect for both boys and girls that can be leveraged relatively cheaply by simply reaping positive externalities from deliberate assignment of children in order to achieve Pareto superior gender compositions of children in classes at the primary school level, for both boys and girls (Dewan et al., 2024).

We finally submit that our study contributes to literature in various ways. One, our findings contribute to the currently relatively small body of peer gender effect literature on developing countries. Two, we contribute to examination of mechanisms of peer gender effects on cognitive abilities. Our findings point to the existence of distinctly gendered pathways through which peer gender effects impact cognitive abilities of girls versus that of boys. Three, we also contribute to the larger debate on single-sex versus coeducational schools; our results illustrating that the effect of single-sex schools is likely quite different for boys and girls. Four, we contribute to the broad literature studying gender gap in math abilities that has been found to lead to occupational segregation by gender, gender pay gap and the overall negative gap in the economic outcomes of women in comparison to men (Joensen & Nielsen, 2016; Rakshit & Sahoo, 2024; World Bank, 2012). We find that having girls study in classes with more girls could potentially prevent the gender gap in math abilities from emerging at the primary school level, which previous studies have found is the period when the gap first shows up (Doris et al., 2013; Fryer & Levitt, 2010; LoGerfo et al., 2006; Robinson & Lubienski, 2011).

The rest of the paper is organised as follows. In Section 2, we describe the data used for this study. In Section 3, we delve into our identification and empirical strategies, including conducting validity tests for identification. In Section 4, we present the results of our analysis, and conduct robustness checks and heterogeneity analysis. In Section 5, we discuss our results and explore the mechanisms of peer gender effect. Finally, in Section 6, we conclude our paper.

## 2. Data

In this study, we employ data collected from the Young Lives (YL) study in India in 2006-07 and 2009. Young Lives (YL) is a longitudinal study which follows two cohorts of children in four developing countries of the global South, namely, Ethiopia, India, Peru and Vietnam. In India, the study was conducted in Andhra Pradesh and Telangana states, and collected data from two cohorts of children – 1008 children of the older cohort born between January 1994 and June 1995, aged around 8 years in round 1, and 2011 children of the younger cohort born between January 2001 and June 2002, aged around 1 year in round 1 – in 2002, 2006, 2009/10, 2013 and 2016. In these rounds of the survey, data was collected using household visits from children and their families and communities. In 2010/11, an additional round of data collection was conducted through school visits for a randomly chosen sub-sample of the younger cohort of children from the panel study. We primarily use data from round 3 in 2009/10, and the school survey round conducted in December 2010, for the younger cohort of children aged 8-9 years in December 2010 for our analysis. As the two rounds being used were conducted within a time period spanning almost one academic year, it allows us to use cognitive scores from round 3 as baseline ability measures, while also using round 3 to provide other community, household and child specific variables, not available in the school survey (Singh, 2015). The school survey, on the other hand, provides us with rich information pertaining to school, class and teacher related attributes of the child.

**Cognitive Ability Measures –**

The main outcome variables in our study are cognitive ability measures gathered between December 2010 and March 2011 from the school survey round (SS1). In this school survey round, the Young Lives study surveyed a random sub-set of schools that were then being attended by the YL panel children of the Younger Cohort (YC). Out of all the children of the YC reported to be enrolled in schools in 2009 in round 3, in Andhra Pradesh, 955 children were randomly selected to be surveyed in SS1, across a total of 247 schools.

These children were administered mathematics, English and Telugu tests, consisting of 21, 29 and 32 items each, in December 2010. The maths test was conducted in two sections covering questions testing basic quantitative and numeric concepts, ability to perform simple mathematical operations and questions on addition, subtraction, multiplication and division. The language tests in English and Telugu comprised of sections testing reading, comprehension and writing in the respective languages.

As lagged measures of cognitive ability, we use cognitive ability scores from round 3 of the survey. We use math test in round 3 as baseline mathematical ability measure, and scores on Peabody Picture Vocabulary Test (PPVT – administered in Telugu) as baseline language ability scores as proxy for lagged measure for Telugu and English language abilities. There were several linking items between the math tests in Round 3 and the school survey round. The PPVT version administered in Round 3 was PPVT-III (Dunn and Dunn, 1997), developed originally in English and then revised numerous times, over time (Leon, 2020). This empirical strategy is in line with value added analysis as modelled by Singh (2015) in order to capture the effect of private versus public schools on cognitive and noncognitive outcomes of the children using the Young Lives data. Singh (2015) too, uses math and PPVT scores from Round 3 survey to proxy for lagged ability measures for analysis on cognitive abilities from School Survey 1 round, for Math, and English and Telugu, respectively.

In order to construct both endline and baseline cognitive ability measures, we employ Item Response Theory (IRT), instead of using the raw scores on cognitive ability tests. Aggregating raw scores assigns the same weight to all the items across a test and does not account for the possibility that items capture varying degrees of difficulty, or different information levels even when designed to capture the same latent ability, for example, mathematical and quantitative abilities. IRT models, therefore, generate measures of latent ability that estimate the probability of getting the response to an item correct, for every item for every candidate, thereby producing a summary score for overall skill for every person.

By enabling us to capture true latent/underlying skill that every item on a test is designed to proxy for, IRT scores allow us the advantage of accounting for heterogeneity in item difficulty, as well as computing scores on the same scale when the same tests are repeatedly administered. We employ the three parameter model (3PL) of IRT, and use OpenIRT command suite written by Tristan Zajonc to generate maximum likelihood estimates for different cognitive ability measures (Das & Zajonc, 2010). The item response function links underlying ability to the probability of getting the answer to the item correct. The basis of 3PL IRT is modelling the behaviour of each test item in order to remove all differences from the score; these three parameters include the item difficulty, item's ability to discriminate between two test-takers of varying abilities, and the probability of the test-taker guessing the right answer.

A unique feature of the YL study is that the cognitive tests administered in the YL are by far the most comprehensive in terms of capturing variation across the skill and learning spectrum among Indian students as opposed to the other comparable tests conducted on the sample of Indian children including ASER by Pratham and the Indian Human Development Survey (IHDS) 2005 that provide only basic measures (Singh, 2015).

Fig. 1 plots a Kernel density distribution of standardized IRT math and PPVT scores from Round 3 for boys and girls. A Kolmogorov-Smirnov test reveals that the distributions for boys and girls are neither significantly different in the baseline for math nor for PPVT, as we fail to reject the hypothesis that the two distributions are equal for math or PPVT.

**Peer gender composition variable –**

The main independent variable in our study is the peer gender composition that we capture as the proportion of girls in the same classroom. This variable is obtained from the School Survey done in December 2010. Students are typically assigned to classes in the beginning of the school year in India and continue in the same classroom throughout the academic year. In most schools in India, students remain in the same classroom for the whole day while teachers come to take classes, and they also participate in various activities together. We operationalize student's peer gender composition as the proportion of girls in the same classroom as the student, excluding the $i^{th}$ student (Gong et al., 2019).

46.5% of students in the sample were females and the average proportion of females in the classroom is 48.85%; the average proportion of females is 51% in the female sample and 47% in the male sample, as shown in Table 1. In Figure 1, Panel A plots the original distribution of leave-one-out proportion of females in the sample, and Panel B graphically represents the conditional distribution of the leave-one-out proportion of females across the sample. The figure shows that there is sufficient variation in the classroom gender composition variable, across classes. Additionally, we find that the $1 - R^2$ derived from regressing classroom female proportion on all controls, including school fixed effects equals 0.746, which indicates sufficiently high residual variation across gender composition across classes (Gong et al., 2019).

**Data from school survey –**

Besides school-end measures of cognitive ability in Mathematics, English and Telugu, and the peer gender composition of the classrooms, the school survey provides us with a rich battery of school, class, teacher and student related variables.

In the school survey, mainly three types of schools were included, namely, public (government) schools, private aided schools and private unaided schools. Nearly 62.6% of the schools were government, while 35.5% were private unaided schools. A meagre 1.9% were private aided schools. This is in line with the overall state trends of Andhra Pradesh in 2011, as reported in ASER (2011) which shows that 62.3% of all children aged between 6-14 years were enrolled in government schools, 34.7% were enrolled in private schools, while 0.3% were enrolled in other institutions. Table 1, Panel C shows that girls are more likely to study in government schools, while boys are more likely to study in private unaided schools as the t-test for the difference between the gender is significant at 1%.

We also include a set of class-specific variables gathered from the school survey. While the participants in our sample are aged 98-118 months with a standard deviation of 3.8 months from a mean of 107.3 months, not all students are in the same grade in the school. The school survey was meant to capture students at the cusp of primary school in grade four or five. However, only 38.4% and 18.3% students from this sample are in grades four and five, respectively. 43% students are in fact in lower grades. This is in line with ASER's findings and the general state of primary school education in India where students are often enrolled in lower grades even at older ages. The average representative "girl" studies in 3.7th class, while the average representative "boy" studies in 3.4th class, and the difference is statistically significant.

The school survey also reports the average performance of the class in Math, English and Telugu, as well as the strength of the class. These variables are summarized in Table 1. The average student to teacher ratio in our sample is 25 for girls and 26 for boys, with no statistical difference between the two.

Finally, the school survey reports on several attributes of the math teacher including her gender, caste, religion, qualifications, nature of employment as temporary or permanent, and four four-point items related to the teacher's gender attitude. We use the "swindex" command suite in Stata written by Schwab et al. (2020) that creates an index of teacher's attitude by standardizing the inverse-covariance weighted average of teachers' self-reported subjective responses to the gender related questions . The weighted average in place of simple average allows us to capture the extent of teacher bias on the probability of one respondent to reply "agrees (disagrees)" to the gender question, but also on the variation in one respondent's "agrees" to another's "strongly agrees". A higher score on the gender attitude index indicates that the teacher has a more gender positive attitude. Girls in our sample are statistically more likely to have teachers with permanent employment contract. We conjecture that this might be a function of girls being more likely to study in government schools. These math teacher attributes are summarized in Table 1, Panel D. The survey does not capture corresponding attributes of English and Telugu teachers.

**Data collected through household survey –**

Besides round 3 cognitive scores, an extensive set of child, parent and household variables, pertaining to the predetermined characteristics of the child, are also captured from round 3 data.

The region variable captures the region in (united)Andhra Pradesh where the household is situated, namely, Coastal Andhra, Rayalseema or Telangana. There are systematic differences and disparities in the geographical and socio-economic conditions of these regions. In Coastal Andhra, known as the "Granary of South India", in 2004-05, while only two out of nine districts were classified as "poor" based on the cutoff of 1.25 USD per capita per day, in Telangana, 10% of the rural population in five out of ten districts lived in poverty, and in Rayalseema, three out of four of its districts were classified poor (Reddy & Bantilan, 2012). Socially too, there are major demographic differences between the three regions. While Coastal Andhra has the highest population density, Telangana has the highest share of scheduled castes and scheduled tribes, whose socio-economic status is markedly inferior to that of the upper castes, and Rayalseema, while having similar levels of human capital indicators as Coastal Andhra, is a region of very low agricultural productivity, and therefore, general levels of prosperity (Reddy & Bantilan, 2012). In our sample, nearly one third girls and boys come from each of the three regions (Table 1, Panel B).

The Round 3 survey also provides information on whether the household resides in an urban or a rural area. Ours is a rural-heavy sample with 80% of children coming from households located in rural areas. The survey indicates that 25% of mothers were only primary school educated, 27% were only secondary school, 2% had had higher education of some kind of the other. The survey also provides information on the wealth of the household using a wealth index, constructed using three sub-indices – housing quality, consumer durables and access to services. We find no difference between girls' and boys' distributions on wealth quartiles.

In terms of the other predetermined characteristics of the child, we find that 90 % of children in our sample are Hindu, 5% are Christian and 5% are Muslim. The majority of the individuals, nearly 50%, in our sample come from the backward castes, and about 30% are SC and ST. Nearly 84% of the children report that their mother tongue is Telugu, the local language of the Andhra Pradesh region. The average number of siblings that an individual in the sample has is 1.63, and 67% are firstborns, that is, the eldest born progenies to the same mother. All of these individual and household level variables are summarised in Table 1. We find no differences between the girls' and boys' distribution on any of these attributes, as can be seen from t-statistics reported in column 3.

## 3. Identification and Empirical Strategy

### 3.1. Identification

In our study, we analyse survey data from primary schools in Andhra Pradesh and Telangana, to estimate the effect of peer gender composition on students' cognitive outcomes.

Manski (1993) raised three concerns with respect to identification of peer effects: one, called the "reflection problem" which arises from endogenous peer effects as peer group members' outcomes are simultaneously determined making it difficult to isolate the effect of the group on the individual. The second issue is that of "endogenous sorting into groups" where identification becomes an issue because of the tendency of similarly endowed individuals, or individuals with similar attributes to aggregate into same groups. And the third issue that he highlights is that of "correlated unobservables" which arises because of the possibility of attributing the effect of shared external factors that comes with being part of the same group, such as institutional and environmental factors, to the effect of peers.

Our identification does not suffer from the reflection problem because we are estimating exogenous peer effects, which is the effect of the predetermined characteristics of the groups on the outcomes of the individuals. Yet, we recognize three potential threats to our identification. One, parents might still choose schools with a desirable gender composition in mind, leading to endogenous selection, two, omitted variable bias from child level unobservable characteristics that might be correlated with both peer gender configuration and the child's cognitive outcomes, and three, shared environment factors that are correlated with both peer gender composition and child's cognitive outcomes (Dewan et al, 2024; Jain & Kapoor, 2015) might confound peer gender effect.

In order to draw implications for policy and analysing the effect of the gender configuration of a child's peer group at primary school level (or below), it is desirable to derive effects that are not driven by various kinds of other factors including those driven by differences in child, household, teacher, class and school related characteristics (Doris et al., 2013). In this section, therefore, we address these three issues to identification – first, the problem of endogenous sorting of students across schools, second, omitted variable bias in the ability production function, and third, the problem of correlated confounding effects.

#### 3.1.1. Discussion on Identification in the context of Indian YL schools, or The Problem of Endogenous Sorting –

The variation in the gender composition of the child's peerage can stem from two sources – one, the natural gender variation in the demographic composition of the child's cohort (Anelli & Peri, 2019; Dewan et al., 20124; Hoxby, 2000), and two, assignment of the child to a class within her school (Anelli & Peri, 2019). The first can naturally be considered exogenous to the child (Anelli & Peri, 2019). The second, while not entirely perfectly determinable a priori by parents and students on account of class assignment within schools being de facto random (Dewan et al., 2024; Eisenkopf et al., 2015), depends at least in part on the choice of the school, which may not be random (Gong et al., 2019). Thus, it is of merit to look at the various determinants of primary school choice for Indian parents, and how that choice plays into the dynamics of the gender composition of a child's peerage.

In order to understand the determinants of school choice, we primarily employ insights directly from qualitative research carried out by the Young Lives (YL) survey based on in-depth interviews with parents of children aged between 9 and 10 years in mid-2011; purposively sampled from the sample of Younger Cohort (YC) children from the YL who were surveyed in the School Survey Round 1 in December, 2010, supplemented by other related studies in the Indian context. This is a sub-set of the sample we use in this study. Interviews were conducted with thirty parents across three YL sites in mid-2011 to explore factors affecting the decision-making process related to choice of school and school change decisions (Morrow & Wilson, 2014).

As proffered by the theory of Rational Choice, "school choice" decision is an elaborate one that rests on multiple considerations on the part of the household including the availability of schools, parents' perception of the school's quality, accessibility, and affordability (James & Woodhead, 2014). In line with this, the YL qualitative study of the parents of the YC highlights the following non-exclusive, often overlapping parameters as factors of central consideration for school choice. One, school quality. While the construct of school quality itself is based on several of the following factors, one important consideration is the perception of teacher presence, attention, class size per teacher, and school amenities. YL parents reported that they perceived private schools to be of better quality than public schools, based on these parameters. The second important factor is securing their children an English education as it has a high perceived, aspirational and actual value in terms of labour market returns (Morrow & Wilson, 2014). Third factor is the choice of source of information employed in order to determine the school quality. Individuals with different resources – physical, social and intellectual – privilege different sources of information in order to ascertain the quality of schools. Four, choice of child in the household to prioritize based primarily on child's sex, birth order, and to some extent, child's academic/market potential. In a household with multiple children and considerable resource constraints, instead of sending all their children to "good" schools (based on parents' perceptions and their private criteria for what constitutes "good"), parents may choose to send those children to better establishments who they perceive to have the most potential or would benefit the most from it, instead of "wasting money" on those who they think won't study as well. These considerations inevitably intersect with the sex and birth-order of the child. Eldest sons are likely to be preferred candidates for a private, English education, as opposed to younger siblings, or girl children in the household. Which brings us to another important factor, especially in resource constrained households, that is, financial considerations and affordability. While a private, English education is desirable, not all households can afford to send all their children to these schools. Five, the study highlights certain other factors that may eclipse quality as the most important factor. In settings of limited school supply, distance to school supersedes school quality.

While in the YL qualitative study on YC parents the researchers do not report any preference of parents with regards to gender profile of the school, numerous studies on parents' preference for single-sex schools have contended that there seems to be a pervasive view amongst parents that students, especially females, tend to learn better in the company of peers of the same sex (Doris et al., 2013). Thus, parents may choose a particular school in response to the gender composition of the female peers (Black et al, 2013; Hu, 2015), and in selecting single-sex schools, the most significant factor for parents is their belief in peer gender effects (Sohn, 2016).

Based on these insights, we draw three conclusions. One, in the YL study sample from Andhra Pradesh, India, gender composition of the school does not seem to be a dominant basis for school selection amongst households in the first place. In fact, only 29 children from a sample of 955 study in single-sex schools, 7 in boys-only and 22 in girls-only schools. This provides one basis for assuming exogenous variation in classroom peer gender composition in our study because of no revealed preference for single-sex schools.

Two, while there are numerous factors to be considered while making the school choice, the actual set of parameters over which Indian parents can choose is the following – public versus private school, English versus non-English schools, geographically closer versus schools situated farther away, single sex versus coeducational schools, and high fee versus low fee schools. And these parameters themselves must also be prioritized as they are most often available as indivisible bundles of choices. Thus, not only are these considerations not entirely disjointed from each other, parents' or households' ability to elect from amongst various available school choices is limited to choosing along these individual facets of schools, making perfect sorting of students of similar attributes into same schools unlikely. Three, conditional on choosing the school, neither parents nor students have any say in classroom assignment, rendering it de facto random (Eisenkopf et al., 2015).

Our identification relies on exogeneity of assignment of students to classes, generating random variation in the peer group makeup (Sacerdote, 2011). We show that while parents and students non-randomly select schools, they have no influence on the classroom assignment, making it, in effect, random. We, therefore, solve the problem of endogenous sorting into schools by including school fixed effects (Dewan et al., 2024, Hu, 2015; Sacerdote, 2011). Adding school fixed effects controls for all school-level observable and unobservable factors that might result in endogenous sorting while allowing us to exploit the idiosyncratic variation in gender composition of students across classrooms. A number of studies on peer effects have exploited the exogenous variation in classroom assignment, conditional on school choice, as identification strategy (Anelli & Peri, 2019; Carrell et al., 2009; Dewan et al., 2024; Gong et al., 2019; Hoxby, 2000; Hoxby & Weingarth, 2005; Hu, 2015; Lavy and Schlosser, 2011; Sacerdote, 2001; Shue, 2013; Zimmerman, 2003; inter alia).

#### 3.1.2. The Problem of Omitted Variable Bias –

Despite controlling for school fixed effects, there might still be other child level factors, observable and unobservable, that might lead to omitted variable bias. In order to mitigate this issue, we include child and household level factors as controls, while employing the value-added model of cognitive abilities.

Background characteristics of the child are important determinants of the student's cognitive ability including student's age, grade, geographical region they belong to, whether they reside in urban or rural communities, their religion, caste, mother tongue, etc. As we discussed earlier, when household resources are limited and gender norms are strong, birth order as well as the sex of the child has a bearing on the school they go to (Morrow & Wilson, 2014). Mother's education has been shown to be correlated with child's cognitive outcomes (Singh, 2015). Thus, we include all of the above and a measure of household wealth to our model. As the likelihood of growing up in mixed gender households is higher for children that have a higher number of siblings (Doris et al., 2013), we also include child's number of siblings (as well as siblings squared) as controls (Brenoe & Zolitz, 2020).

As mentioned earlier, from the qualitative in-depth interviews of the parents of the YC of the YL children being studied here, we find that parents' sometimes privilege other factors besides the quality of school while choosing schools for their children such as the physical distance of the school from their residence (Marrow & Wilson, 2014). This creates a selection problem on two related counts. One, often better quality schools tend to be located in more urban, more prosperous, more developed localities, which are known correlates of child's cognitive and noncognitive abilities (Almlund et al., 2011) *add more references*, creating a systematic geographical basis for endogenous selection of school choice to child's abilities. Two, the availability, or lack thereof, of good quality schools in a household's vicinity within whatever a household deems a "comfortable enough physical distance", diminishes the child's likelihood of accessing "superior" quality schools by limiting the pool of schools available to them. However, this accessibility is not only a function of the distance of the household from the school that households choose to send the child to, but also of resources, such as the means of travel the household can afford to access farther located but deemed to be better quality schools. This also creates a systematic selection problem. Therefore, we control for the time taken to reach school which happens to be a function of both distance and mode of transport available to the child to travel to school thereby accounting, to some extent, for this two-fold endogeneity.

Finally, we choose the dynamic ordinary least square (DOLS) model for value-added analyses, which Guarino et al. (2011) observe was the most robust estimator across various scenarios, and by adding lag of cognitive ability as a measure of baseline ability, controls for "grouping and assignment mechanisms" (Singh, 2015). By including baseline ability measures, we control for individual-level variation across student quality. Our identification strategy is similar to Dewan et al. (2024) and we improve upon Hu's (2015) specifications by controlling for child's lagged ability, accounting for the contributions of all previous inputs, shocks and variations in unobservable individual factors, that might cause peer gender composition to be correlated with the error term.

The identifying assumption in our study, therefore, is that conditional on covariates, lagged ability and school fixed effects, the variation in peer gender composition is uncorrelated with the error term.

#### 3.1.3. The Problem of Confounding Peer Effect with Correlated Shared Environment Effects –

The problem of correlated effects (Manski, 1993; Sacerdote, 2011) arises from the possibility of erroneously attributing shared environment effects to peer influence. In our estimation, it is partly checked by including school fixed effects as that controls for school-level characteristics that might impact both the classroom gender mix as well as students' cognitive outcomes. However, there are other possible confounders.

Peer effect could as easily be a result of shared classroom environment and teacher effects. Dee (2007) shows that having a female teacher adversely affects math outcomes of both boys and girls, whereas Carrell et al. (2010) find that top performing girls benefit from having a female teacher in math and science. On the other hand, Winters et al. (2013) did not find any effect of teacher gender in the early years of schooling but a significant positive effect on grades in middle and high school. Cho (2012), too, finds no effect of teacher gender on grades across fifteen OECD countries. Rakshit and Sahoo (2024) show that girls are negatively affected by teachers who harbour gender biases against females even when girls and boys study under the same teacher. This is in line with Carlana (2019) and Alan et al. (2018) who also find that more gender-biased teachers cause a negative impact on child's cognitive abilities, and more qualified teachers with higher recorded cognitive abilities are less likely to sport negative gender biases. In our study, we have teacher's gender, qualification, age, caste and gender attitude index, as explained in Section 2, for the mathematics teacher. Thus, for the math analyses we are able to control for teacher attributes that might impact the student outcomes in the class, and show that the observed peer gender effects are not driven by teacher effects. However, even though we don't have these variables for English and Telugu teachers, we argue it out in detail in Section 6 and show that even in these cases, peer gender effects in English and Telugu most likely cannot solely be explained away by teacher effects.

Classroom level factors such as student strength in the classroom and classroom gender norms and practices can also impact cognitive outcomes of children, once again leading to potential confounding effect on peer gender effect estimates. Borbely et al. (2022) highlights how having more students in the classroom can effectively diminish teacher attention per student, and potentially make the role of peers even more salient. Morrow and Wilson (2014) report in the qualitative YL study that parents perceive students to teacher ratio as a measure of teacher effectiveness. Further, classrooms with stricter gender norms can potentially limit inter-gender interactions both in and outside classes, confounding the peer gender effect. Therefore, we control for student-teacher ratio of the classroom and include a dummy that takes the value one if students of opposite sex are mandated to sit separately in the classroom, and zero if there is no such mandate preventing inter-gender mixing inside classrooms.

Another major factor that can easily be confounded for peer gender effects is the ability spillover from peers; one of the concerns highlighted in the earlier section, as discussed by Manski (1993). What we capture as peer gender effects can plausibly be the effect of having high ability peers. Therefore, we include average peer performance of the class in Math, English and Telugu. These measures are good proxies for the average quality of peers in the class (Rakshit & Sahoo, 2024), as well as the effectiveness of teaching in the classroom. Controlling for the average peer performance allows us to control for the endogenous impact of peer group on the child via ability spillover, thereby ensuring that we only capture the exogenous impact of the peer gender mix on the child (Gong et al., 2019; Sohn, 2016). In Section 6, we argue in greater detail as to how and why our results are most likely not confounded by ability spillover either.

While we already include school fixed effects, we also include a few observable school characteristics from data. We factor school type based on whether the school is public, private unaided or private aided . As already discussed, parents choose between private and public, and private low-fee or aided and unaided for a variety of reasons ranging from perceptions about quality of education, teacher, facilities available, availability of instruction in English, etc. We also directly control for whether the school offers instruction in English or not. In backward areas especially, given the egregious lack of resources, funds and facilities, many schools lack basic facilities like separate toilets for girls, or even any functioning toilets at all. Separate and functioning toilets for girls seems to have a significant bearing on the parents' choice of school in such areas, especially for girls (Morrow & Wilson, 2014). In Table 1, we see that girls are at least 1.7% statistically significantly more likely to go to schools with gender segregated toilets. Therefore, we also include an indicator for gender segregated toilets for girls.

We address the problem of correlated effects because of a common, shared environment by including various school, class and teacher level controls. In section 6, we deal in even more detail with the possibility of incorrectly attributing the effects of shared environment to the effect of peer gender composition.

### 3.2. Empirical specification –

Based on this, our main estimating equation for investigating the effect of peer gender composition on the child's ability, is a linear-in-means value-added model of cognitive abilities as specified in equations (1) through (8), in line with several above mentioned studies in this literature. We estimate the following models across math, English and Telugu, separately for males and females, following Hu (2015).

$$Y_{ics}^{t} = \beta_0 + \beta_1 FemalePeer_{ics} + \beta_2 Y_{ics}^{t-1} + \varepsilon_{ics} \quad (1)$$

$$+ \beta_3 X_{ics} \quad (2)$$

$$+ \eta_s \quad (3)$$

$$+ \beta_4 W_{ics} \quad (4)$$

$$+ \beta_5 C_{ics} \quad (5)$$

$$+ \beta_6 S_{ics} \quad (6)$$

$$+ \beta_7 Z_{ics} \quad (7)$$

$$Y_{ics}^{t} = \beta_0 + \beta_1 FemalePeer_{ics} + \beta_2 Y_{ics}^{t-1} + \beta_3 X_{ics} + \beta_4 W_{ics} + \beta_5 C_{ics} + \beta_7 Z_{ics} + \eta_s + \varepsilon_{ics} \quad (8)$$

where $Y_{ics}^{t}$ denotes the standardized cognitive ability scores for Math, English and Telugu estimated using IRT for student $i$ in class $c$ in school $s$, measured from the school survey at time $t$. $FemalePeer_{ics}$ is the measure of peer gender composition operationalized as the proportion of females in the classroom, excluding self. $Y_{ics}^{t-1}$ captures the baseline cognitive and noncognitive abilities of the child.

Equations (2)-(7) sequentially include vectors of various kinds of controls.

$X_{ics}$ is a vector of various student and household level controls for the student i. $\eta_s$ captures school fixed effects. The vector $W_{ics}$ effectively controls for the average quality of peers in the class (Rakshit & Sahoo, 2024), and the effectiveness of teaching in the classroom, as it contains the average performance of the peers across Math, English and Telugu. The vector $C_{ics}$ controls for shared class level characteristics. $S_{ics}$ vector contains various school characteristics that might influence the choice of school and might also be correlated with the child's abilities. The vector $Z_{ics}$ controls for a set of math teacher characteristics.

For all the models, the standard errors are clustered at the school-class level in order to account for heteroskedasticity and possible serial correlation in the outcomes of students in the same school and class.

### 3.3. Validity Tests –

In this section, we assess the empirical validity of our research design by examining certain conditions – balance tests on predetermined characteristics, noncompliance behaviour, and attrition (Sohn, 2016).

#### 3.3.1. Balance tests on predetermined characteristics –

Our identifying assumption for estimating peer gender effects is one of conditional exogeneity, that is, conditional on various covariates, school fixed effects and lagged ability, the residual variation in the peer gender composition of the classroom is idiosyncratic and uncorrelated with the error term. In section 4.1, we discuss that it is unlikely that there is any sorting on the basis of peer gender configuration of the classroom in our sample. We now test if that assumption of as-good-as-random assignment holds.

We consider various student and household characteristics in our data and regress each of them on the proportion of females in the classroom, our treatment variable. Table 2 shows the results of the balance tests. Columns (1) and (2) show the balance test results for the entire sample, while columns (3) and (4), and (5) and (6) report balance across male and female sub-samples, respectively. Columns (1), (3) and (5) report unconditional estimates, while columns (2), (4) and (6) report estimates conditional on school fixed effects.

The unconditional estimates show that the predetermined characteristics of the students vary somewhat with the proportion of females in the class, such as student's baseline noncognitive abilities, mother and father having high education, and the household belonging to the lowest wealth quartile, albeit with coefficients of negligibly small magnitude. Controlling for school fixed effects makes these statistically insignificant. The few predetermined covariates that are statistically significant upon the inclusion of school fixed effects, including baseline PPVT, or receptive vocabulary, score, belonging from Rayalseema and residing in an urban area, have coefficients that are very small in magnitude and are therefore, not economically significant. For example, the coefficient of belonging from Rayalseema, the factor with the highest point estimate, implies that a ten percentage point increase in the proportion of females in class corresponds to a 1.3 percent lower likelihood of the child belonging from Rayalseema (1 percent higher for males and 1.7 percent lower for females).

These tests sufficiently indicate that there are negligible differences in the predetermined characteristics of the child across peer gender composition, confirming the validity of our identifying assumptions.

#### 3.3.2. Non-compliance, or the issue of mid-cycle school migration –

As we capture the classroom proportion of females and the cognitive outcomes of the child at the same time, from the School Survey Round 1 in December of 2010, the validity of our identification does not require us to make strong assumptions pertaining to students migrating out of school subsequent to classroom allocations at the beginning of the school year. Noncompliance to treatment allocation in that sense relates to students switching schools in response to the knowledge of the exact gender composition of their classes becoming available to them. But because of the nature of data collection, even mid-cycle school migration cannot possibly be endogenous in our design.

Switching of schools mid-year does, however, allow for at least one problem to persist, that is, it generates variations in treatment dosages across the sample. Not only can dosage change because of mid-cycle migration of the index child, but also due to (in or out) migration of classmates at any point during our period of analysis. While for the latter to be a source of concern, the size of classmates' migration must be significantly high for the treatment variable to sufficiently deviate, arguably, the first is indeed a plausible source of concern. As such, given that the YL school survey does not record information about migration of the children or their peers during the particular academic cycle, our analysis is limited by data feasibility in controlling for differences in dosages of exposure to the recorded gender composition of the YL children's classrooms.

#### 3.3.3. Attrition –

There are several missing values across the outcome variables, the treatment variable and the other covariates specified in our estimation equation(s). We, therefore, need to address the problem of sample attrition and check if there is any correlation between the likelihood of missing values and the treatment variable (Dewan et al., 2024; Doris et al., 2013; Rakshit & Sahoo, 2024). We create an attrition dummy that takes the value one for any missing variables and regress it on the proportion of females in the classroom, and on school fixed effects. As we can see from Table 3, the coefficients of proportion of females in the classroom are all very small in magnitude and statistically insignificant. This indicates that our results are not driven by attrition in the sample, and it does not bias our results presented in the following section.

# 4. Results

## 4.1. Main Results – Peer gender effects on Cognitive Abilities

We now estimate the treatment effect of classroom peer gender composition on students' cognitive abilities, separately for boys and girls, as in Hu (2015), based on specifications enlisted in equations (1) through (8). As shown in Table 4, we report the results of an unconditional dynamic ordinary least square (DOLS) model in the first column. In the subsequent columns, we report the results of sequentially adding various sets of controls to this base model – student and household controls in column (2), school fixed effects in column (3), average peer ability controls in column (4), class controls in column (5), school characteristic controls in column (6) and math teacher controls in column (7). In column (8), we run the regression corresponding to equation (8) specified in Section 3. Panel A reports results for females and Panel B for males.

The results show that having more female peers has a statistically significant positive causal effect on the cognitive abilities of the students, on an average, and that these results are distinctly heterogenous by gender. Panel A of Table 4 shows that, on an average, having female peers statistically significantly improves mathematical and quantitative abilities, as well as English language abilities of females in the same class, for some specifications, but has no impact on their Telugu abilities. In order to appreciate the economic significance of these estimated effects, we use the more conservative estimates from equation (7) for math abilities, reported in column (7) in Table 4, where we control for lagged student ability, student-household characteristics, school-class-teacher characteristics, peer ability as well as school fixed effects. The coefficient of 1.04, significant at $p < 0.05$, means that a ten percentage point increase in the proportion of females in the classroom would cause the mathematical and quantitative abilities of females to improve on an average, by 10.4% of a standard deviation. These estimates are in line with the results obtained by Gong et al. (2019) where they also observe an effect size of 1.299 for mathematical abilities in their sample, corresponding to increase in the proportion of females. Their coefficient for females equals 0.148 and is statistically significant.

On the other hand, Panel B of Table 4 shows that, on an average, an increase in the proportion of female peers in the classroom does not have a robust causal impact on the mathematical abilities of the males in the classroom, even though adding math teacher controls in specifications (7) and (8) suggests that there might be some improvement to the tune of about 10 percent of a standard deviation in math abilities in response to a ten percentage point increase in the proportion of female peers. Males, however, undoubtedly experience a robust causal improvement in their language skills in English, as a result of having more female peers. We find that the coefficient of proportion of females is statistically significant and positive for English, for males, across various specifications. Column (6) coefficients mean that a ten percentage point increase in the proportion of females in the classroom causes English abilities of the males in the classroom to improve by almost 14 percent of a standard deviation, on an average. Once again, we see no impact on Telugu abilities for males either.

### 4.2. Robustness Analyses –

We now present a number of robustness checks of our main results on Math, English and Telugu abilities of primary school children as a result of their classroom gender composition. In section 3.3.3, we have already demonstrated that attrition from accounting for missing values across various covariates does not bias our analyses.

#### 4.2.1. Restricted Sample –

In our analysis, we construct the peer gender composition measure as a continuous variable measured as the fraction of females in the classroom that goes from 0 to 1. As observed earlier, parents often choose schools on the basis of whether the schools are single-sex or coeducational (Black et al, 2013; Hu, 2015). However, as we have already shown in the previous section, conditional on choosing schools, the actual proportion of females in the classroom is random. However, this conditional randomness of proportion of girls in the classroom does not apply to single-sex schools. An admittedly more accurate identification would be to say that conditional on choosing coeducational schools, the proportion of females in the classroom may be considered as-good-as-random. But because we include the entire spectrum of the fraction of females in the classroom, including 0 and 1, which indicates boys-only classes, and girls-only classes, respectively, we are not able to take into account the non-arbitrariness of female classroom proportion taking the value 0 and 1. Of course, including school fixed effects checks that to some extent. However, in order to check if our results are unduly impacted by the presence of single-sex schools, we also run all our analyses on the restricted sample including only coeducational schools. As mentioned earlier, in our sample, 29 students go to single-sex schools out of which 7 males go to boys-only and 22 females go to girls-only establishments.

Using the restricted sample, we find that results from these analyses closely track our main results. However, almost all significant results improve in magnitude. The most significant difference in analysis from Table 4 to Table 5 occurs for peer gender effects on English abilities for girls. With the restricted sample, we find that upon adding school fixed effects, all our coefficients on peer gender effect become statistically significant at 5% level, implying that our initial estimates for girls' sample were in fact sensitive to the presence of girls-only schools. Upon excluding them, we find that because of an increase of ten percentage points in the proportion of females in the classroom, girls experience an 11.5 percent of a standard deviation improvement in their English abilities. We find that these results are more reliable and robust. Therefore, for females, we think that the effect size would be closer to the estimates observed from the restricted coeducational schools-only sample. The effect size for math abilities for girls also improves. We find that a ten percentage point increase in proportion of girls in the classroom now causes an almost 11.6 percent of a standard deviation improvement in math abilities for girls.

The results for the male sample are hardly changed at all. We still find positive but sensitive improvements in math abilities of males, and a positive and robust improvement in their English abilities to the tune of 14 percent of a standard deviation because of a ten percentage point increase in the proportion of females in the class. We still find no effect on Telugu abilities for either boys or girls.

We, therefore, conclude that accounting for the possibility of systematic differences between single-sex and coeducational schools actually improves our estimates of peer gender effect on girls, while remain much the same for boys. Thus, our estimates of peer gender effects are quite robust.

#### 4.2.2. Randomized Inference –

In order to not rely on the asymptotic properties of our estimators alone in order to draw inferences and given the sample sizes in our analyses (Canaan & Mouganie, 2023; Heß, 2017), we also conduct randomization inference based on our identifying assumption of conditional randomization of peer gender composition. We compute p-values based on randomization inference for all estimates in our main as well as the restricted sample results. We report the same in square brackets in all results (Tables 4 and 5).

These p-values are estimated based on simulations conducted by re-randomizing students to varying peer gender compositions, 1000 times each, based on our identification assumption that conditional on covariates, lagged ability and school fixed effects, peer gender composition is as-good-as-random. The randomization inference p-value represents the proportion of times the estimated treatment effects come out to be larger than the observed treatment effect from the data. If the data captures true treatment effect, then we would expect the empirical coefficients of proportions of females in the classroom from re-randomizations to rarely exceed actual estimates. Thus, in a way, the randomization inference p-values represent the probability that an equal or similar magnitude treatment effect would have been observed under other hypothetical realizations of the chosen method of randomization, and the smaller the p-value, the less likely it is that we would have observed the treatment effect observed under a different random assignment of the treatment.

From the results reported in Tables 4 and 5, we observe that for all our significant coefficients, the randomized inference p-values come out to be very small; all of them less than 0.05 (a threshold found adequate to look at by Gong et al., 2019), in line with results from our actual data. That is, less than 5% of the randomized estimates of $\beta_1$ (coefficient of proportion of females in the classroom) come out to be larger than the observed values of $\beta_1$. Thus, our results can be interpreted to reflect the treatment effect, and are not likely to be caused because of the sampling strategy.

#### 4.2.3. Potential Bias from Unobservables – Bound Analysis

In our main DOLS specification, we control for a large set of observable explanatory variables, including lagged ability and school fixed effects. Nevertheless, if there are any other unobserved, unaccounted for factors which affect both peer gender composition and the cognitive abilities of the child, then our estimates could be biased. We, therefore, investigate the extent to which omitted variables could be biasing our estimates following the methodology proposed by Altonji et al. (2005), and extended by Oster (2013; 2019). This methodology rests on the idea that selection on observables could meaningfully guide assessments about selections on unobservables.

Using Oster (2013), it is possible to comment on the effect of unobservables by focussing on two parameters: $\delta$ and $R^2_{max}$. $R^2_{max}$ is the theoretical $R^2$ of a hypothetical regression which includes the complete set of controls including all the observables and unobservable determinants of cognitive abilities of the child. Making an assumption on the value of $R^2_{max}$ is more a matter of discretion based on theory and empirics, but Oster (2019) suggests that $R^2_{max} = \min(\pi. R^2_{controlled}, 1)$, should be used, where, $R^2_{controlled}$ is the $R^2$ from the regression including all the observable controls, and $\pi$ is a multiplier. Oster (2019) finds, based on simulations of data from papers using randomized data, that $\pi = 1.3$ should be used as a cutoff value. In our study, observing the $R^2s$ of the various cognitive abilities from all the regressions (Tables 4 & 5), we find quite large $R^2$ values to the tune of 70-80%. In our case, 1 lies in between the range of $(1.3 * R^2_{controlled} = 0.7, 1.3 * R^2_{controlled} = 0.8)$. Therefore, for this section, we take an $R^2_{max}$ of 1.

The next parameter of interest is $\delta$, or the degree of proportionality parameter. $\delta = 1$ indicates that we expect that the selection on unobservables is as high as the degree of selection on observables. In our analyses, we don't expect that the degree of selection on unobservables would equal selection on observables. Thus, we calculate $\beta$s corresponding to the $\delta$ equal to 1, 0.5 and 0.25. Columns (3) through (5) of Table 6 report the same. We find that $\beta$s corresponding to $\delta = 1$ (column 3) are all in the same direction as $\beta$s estimated from controlled regressions (column 2). This indicates that our estimates are robust to selection on unobservables. We also find that the average value of $\delta$ for which the adjusted effect sizes (column 5) match the observed effect sizes (column 2) is less than 1; it is 0.25. Oster (2019) observes that this implies that the value of $\delta = 1$ is just a bound.

Oster (2013) further recommends that, as a robustness check, investigating the value of $\delta$ that nullifies the $\beta$ is a useful test. The idea behind this analysis is to allow $\delta$ to be unrestricted and take any value, and based on that value, discover the power of selections on unobservables in driving the results, which when accounted for by including the $\delta$ parameter, renders the treatment effect zero. Thus, Oster (2013) recommends using $\delta$ as a bounding argument and reporting that value of $\delta$ which produces a $\beta$ value of zero, given the chosen value of $R^2_{max}$. This implies that, on an average, larger absolute $\delta$ values indicate more robust results. Negative $\delta$ values imply that, for treatment effect to be zero, if we know that observables are positively correlated with cognitive abilities, unobservables must be negatively correlated. We report the $\delta s$ for our specification of choice including all controls, and school fixed effects, in Table 6, column (6), for females and males. We find that the $\delta$ values for all our analyses are large (greater than 1) for all cognitive abilities, and negative for English. These findings indicate that for our estimates to be biased on account of omitted variables to the extent of rendering the observed treatment effect null, the selection on unobservables needs to be quite high.

Results from Table 6 imply that it is highly unlikely that our estimates are biased, and suggest a causal interpretation of the effects of peer gender composition on cognitive abilities of boys and girls. This robustness exercise consequently lends more credibility to our findings that a female-favoured change in the peer gender composition, on an average, causally improves the math and English abilities of girls, and the language skills in English for boys, in the same classroom (Altonji et al., 2005; Oster, 2013; 2019).

### 4.3. Heterogeneity Analyses/ Sub-sample Analyses–

Next we conduct various heterogeneity analyses to see which population of students is driving the peer gender effects for girls and boys.

We split the sample on three parameters – one, between public and private unaided schools, two, between schools that offer instruction in English medium and those that don't, and three, between schools that provide gender segregated toilet facilities for girls, and those that don't. We report the results of our analyses in Table 7.

We find that our results are mainly driven by students that go to public schools (Panel A), study in schools that do not offer instruction in English (Panel B), and schools that have the least infrastructural resources available to cater to children proxied by the availability of gender segregated toilets for girls (Panel C). This implies that peer gender composition externalities are the most significant for the least privileged and resourceful set of students, and therefore, becomes even more pertinent from the perspective of framing policy.

## 5. Discussion

We find that an increase in the proportion of females in the classroom causes, on an average, a significant, positive and sufficiently large improvement in the abilities of the students, for both boys and girls, and that this positive peer gender effect is asymmetric by gender. Females improve in their mathematical, and in English abilities too, when considering only the coeducational schools' restricted sample, while males see an improvement in their English abilities, as a result of an increase in the female peers in their class.

Even though in our empirical specifications enlisted in Section 3.2 we explicitly control for endogenous peer characteristics like average peer ability, and aspects of shared environment such as class and teacher attributes, Sacerdote (2011) points out that it is difficult to separately identify the two as peer background itself also affects peer outcome, and therefore, even for studies that exploit exogenous variation in the background characteristics of peers (including this one, and Dewan et al., 2024, Gong et al., 2019; Hoxby, 2000b; Hoxby & Weingarth, 2005; Hu, 2015; Sacerdote, 2001; Zimmerman, 2003, to name a few), "that does not imply that both coefficients (of average peer ability and average peer background) are separately identified"; even as Sacerdote (2011) further notes that adding school fixed effects sufficiently controls for self-selection into peer groups. This is because there is a possibility that unobserved differences in children's abilities may still exist, and these, along with other attributes of the shared environment including teacher and class related characteristics may impact child's cognitive abilities (Sohn, 2016). Thus, it becomes important to examine if the observed peer gender effects in our study are in fact being conflated with either – one, the effects of peer ability spillover, or, two, the effects of shared environment.

**Why not Ability Spillover**

One concern is that the positive treatment effect observed could be on account of the female peers' ability spillover, especially plausible as literature notes that girls are seen to have an advantage in test taking at primary and middle school levels (Gong et al., 2019). We, therefore, investigate this eventuality and show that it is very unlikely for our results to be driven by ability spillover alone.

First, we provide comparison between the baseline cognitive abilities of the girls and the boys (Gong et al., 2019), by comparing standardized IRT scores based on math and PPVT tests administered in Round 3 of the YL. Fig. 1 plots the distributions of standardized math and

PPVT scores for boys and girls from Round 3. The two distributions track each other quite closely. The Kolmogorov-Smirnov tests also reveal that we are unable to reject the null hypotheses that the two distributions, for girls and boys, are different for both math and PPVT, in the baseline. The gender gaps for both abilities come out to be economically very small (mean gender gap of 0.043 and 0.076 standard deviations for math and PPVT, respectively), and statistically insignificant (with corresponding p-values of 0.747 and 0.266). Thus, we find no statistically significant gender gap in the average math or PPVT scores, at the baseline.

Second, in our specifications, we control for the average academic ability of the peers, as shown in Tables 4 and 5. We find that adding peer ability controls improve all our significant results in magnitude, while they all continue to be statistically significant. This strongly indicates that while accounting for the average ability of the peers improves our estimates of peer gender effect, they are not likely to be sole drivers of the results and cannot solely explain the peer gender effects. Previous studies also touch upon these issues and similarly conclude that ability spillover is unlikely to be able to solely explain all the gains in child's ability on account of their peerage (Dewan et al., 2024; Gong et al., 2019; Hoxby & Weingarth, 2005; Lavy & Schlosser, 2011).

Three, we test balance of average peer ability in math, English and Telugu between boys' and girls' samples and find that, while significant, the magnitude of all the coefficients are very close to zero indicating that the average peer ability is fairly balanced between the two sub-samples. The results for the same are reported in Panel A, Table 8.

Three, we find asymmetric, heterogenous effects by ability. Females benefit in math and English, while males benefit only in English, on an average, on account of a female favouring change in the peer gender composition. Ability spillover from girls would cause the outcomes of both boys and girls exposed to greater numbers of girls in the classroom to improve across all domains of ability, or only across those domains of abilities that females are better at. However, one, given that we find no difference between males and females in the baseline in either math or receptive vocabulary, females don't seem to have any demonstrably greater advantage in either domains of cognitive ability to begin with, and two, our results by the logic of ability spillover would seem to be somewhat counterintuitive because of heterogeneity by abilities. If we assume that the stereotype that males are better than girls at math is true, then both males and females would not be expected improve in math. However, we find that the pattern of peer effects goes against the expected pattern of peer ability spillover effects – girls actually benefit in math, while boys benefit in English, after controlling for peer ability – therefore, in line with the findings of Borbely et al. (2022), Gong et al. (2019), and Lavy and Schlosser (2011), among others, we also conclude that it is highly unlikely that the positive treatment effects observed on cognitive abilities of boys and girls are driven solely by ability spillover from female peers.

**Why not Shared Environment Effects (class or teacher) –**

The next source of concern regarding the peer gender effects that we explore is to check if it possibly stems from characteristics of the shared environment that come from the set of class and teacher characteristics.

The first check we have is to include observable class characteristics, which we control for in various specifications. Including class controls neither changes the magnitude of our estimates very much (all positive significant effects increase slightly in magnitude), nor their statistical significance (Tables 4 & 5).

Next, we test balance of class characteristics across peer gender composition to check for sorting on observable class attributes for male and female samples. The results of the balance tests are reported in Panel B, Table 8. We find that we are unable to reject the null hypotheses that these characteristics vary significantly by peer gender composition, indicating that, based on observable class attributes, it seems that class characteristics cannot solely explain all of the observed peer gender effect.

The case of the teacher effects is somewhat less straight forward. For one, as mentioned earlier, we only have teacher characteristics of the math teacher and, therefore, are not able to control for teacher characteristics for English and Telugu. The math results are robust to the inclusion of teacher characteristics, including the teacher's gender attitude index, and do not change much in magnitude, and not at all in statistical significance, upon the inclusion of teacher controls, as can be seen from Table 4. The balance tests of the math teachers' characteristics across peer gender composition, including school fixed effects, reported in Panel C, Table 8, also show that the math teachers' characteristics are balanced in our sample, indicating no selection on math teachers.

If we assume that these teacher patterns would be similar, if not identical, for the English and Telugu teachers, on an average, we can argue that class and teacher effects are not likely sole drivers of our peer effects, and in fact, possibly attenuate the true peer gender effects. We project estimates of peer gender effects for males and females for English and Telugu using teacher controls of the math teacher as proxies for teacher controls of English and Telugu teachers, based on the stated assumption, to see how our estimates change. We find that our results still remain insignificant for both languages for females when we consider the entire sample. For males, we find a slight increase in magnitude for English (from point estimate of 1.38 to 1.54, significant at 0.01 level each), and the coefficient of peer gender composition becomes significant for Telugu at 10 percent level while increasing slightly in magnitude (from 0.53 to 0.82). With the data available to us, we can conclude that these are only indicative estimates and only go to show that the peer gender effects for both English and Telugu, for males and females, are likely to change very little if we were able to control for English and

Telugu teachers' characteristics in our analyses, and that even teacher effect cannot fully explain away the peer gender effects in our study for Math, English or Telugu abilities.

### 5.1. Plausible Mechanisms

Now that we have been able to show that our results are not conflating the effects of the other distinct but correlated aspects of the peer effect with peer gender effects, we aim to explore the plausible mechanisms that might be causing the observed gender-specific causal heterogeneity by ability of the peer gender composition effect.

Literature proposes several channels through which the average peer gender composition might impact child's cognitive abilities. From our review of the literature, we find that these proposed channels can be comprehensively classified to fall under six different heads – one, classroom disruption, two, social interaction, three, competitive environment, four, teacher behaviour, five, approaches to learning, and six, gender stereotyping.

***Classroom Disruption***

Lazear (2001) put forth that the negative effects of disruptive peers are the most significant in explaining peer effects on ability. Subsequently, Lavy and Schlosser (2007) found large positive effects from increase in the percentage of girls in the classroom and suggested that these effects work largely through reduction in classroom disruptions and classroom violence, and through improvements in inter-participant (student-student and student-teacher) relationships in the classroom (Sacerdote, 2011). Hu (2015) finds that the positive effect of more female peers might be channelled through less disruptive, more disciplined school lives of students. Dewan et al. (2024) use School Survey 2 of the YL to point to the possibility that peer gender effects are a result of general reduction in disruptions, and a general improvement in the classroom environment from having higher proportion of girls. Gong et al. (2019) report improved cognitive and noncognitive abilities and through channel decomposition show that one of the channels is improvement in classrooms' environments. All of these studies are based on the idea that females have been shown to engage in fewer externalizing behaviours (Borbely et al., 2022; Bertrand & Pan, 2013) and as such, all students in the class, boys and girls, would arguably benefit from improved classroom environment due to the presence of more girls.

We are limited by data to explore this channel but one way to do this is by looking at differences across peer gender compositions in teacher reported classroom behaviours of the YL children. The YL School Survey 1 captures teacher's perceptions of YL students' classroom behaviours using a three point scale, ranging from "Not true", to "Somewhat true", to "Certainly true" on nine items, namely – one, child is restless and overactive in the class, two, tends to be rather solitary, three, is generally obedient, four, fights and bullies, five, is generally liked, six, is easily distracted, seven, picked on or bullied by others, eight, has many fears and is easily scared, and nine, has a good attention span and sees tasks through to the end. Following Hu (2015), we regress these items on leave-one-out proportion of girls in the classroom and school fixed effects, after harmonizing all items such that higher scores indicate more positive reported behaviours, but unlike Hu (2015) and others, we do not find any significant results indicating differences on classroom behaviours of students by variations in peer gender composition. This does not, however, imply that we can rule out classroom disruption as a channel, but only that we do not have sufficient evidence of confirm the same.

Next, we construct class activity index of positive and negative time use as proportions of positive and negative student activities in the class based on the surveyors' classroom observations captured in the YL School Survey 1 during a mathematics class. We regress the proportion of students' positive and negative activities in the class on the proportion of girls and school fixed effects, as before, separately for males and females. We find that while largely there is no correlation between the probability of students engaging in positive or negative activities and the peer gender make-up of the class, for the female sample, we find a large negative correlation (coefficient of -0.641 with a standard error of 0.37) between proportion of negative activities engaged by students in the classroom and the proportion of girls, on an average. While once again not conclusive, this provides some evidence, albeit a somewhat weak one, of reduction in classroom disruptions being a viable channel for the observed treatment effects.

***Social interactions***

Social interactions play out between the same as well as opposite sex peers and rationales attributed to changes in social interactions between peers on account of sex have been used to justify both positive and negative same and opposite gender effects. Jackson (2021) suggests that boys in single-sex schools likely perform cognitively better than their male peers from coeducational schools as they experience lesser anxiety about impressing their classmates. Black et al. (2013), Coleman (1968), Doris et al. (2013), Hill (2015), Mael et al. (2005) and Sohn (2016) observe that females might be distracted by the presence of males, and vice versa, and therefore, students experience negative effects of peer gender composition favouring the opposite sex given the presence of dating norms.

On the other hand, increase in the proportion of females in the classroom can reduce the exposure of both males and females to bullying at the hands of other males, thereby, improving the learning outcomes of both males and females (Borbely et al., 2022). Further, it has been suggested that in the zeal for impressing girls, boys exert greater efforts at learning, and devote more hours to homework and tutorials (Gong et al., 2019), and therefore, experience a positive effect of peer gender composition that leads to greater exposure to more females.

We would like to argue that the positive and robust peer gender effects on English abilities of especially males can arguably be attributed to the effects of improved social interactions with the opposite sex. While both boys and girls alter their behaviours in the presence of the opposite sex, the burden of application disproportionately falls to the males as per the prevailing gender norms. This coupled with norms of same-sex and opposite sex exchanges being especially starkly different for males, males benefiting more from socially interacting with more females would plausibly be in English. In India, English is still considered an aspirational language, as opposed to Telugu which is a local language. We can more confidently claim that opposite sex social interaction governed by social norms of inter-sex exchange are driving English results for males, because there is no corresponding improvement in Telugu abilities of either boys or girls, which would have happened had being around more girls simply improved the language skills of their peers.

***Competitive Environment***

The third plausible channel is the changes in the competitive environment wrought by the change in the peer gender compositions because of inherent or learnt differences in responses to competition by gender (Doris et al., 2013). Some studies report that women, under pressure of competition with men, become less confident (Bengtsson et al., 2005; Niederle & Vesterlund, 2011). Further, Gneezy et al. (2003) and Niederle and Vesterlund (2007) demonstrate that males and females exhibit varying degrees of competitive behaviours in same-sex settings; these studies show that it is not that females are less willing or less able to perform well in competitions, but that they do not "compete well in competitions against men" (Doris et al., 2013). Studies by Booth and co-authors (Booth et al., 2014; Booth & Nolen, 2012a; 2012b) provide experimental evidence that females are more likely to engage in competitive behaviours in all-female settings. On the other hand, studies have shown that men become more competitive around other men (Anelli & Peri, 2019).

These studies suggest that an increase in the proportion of girls in a classroom would likely create a more female-friendly environment in class with respect to perceived competitiveness, especially for girls, thereby creating less competition-threatening conditions, in which, females typically tend to thrive.

Our finding that females, on an average, in general cognitively gain across domains when there are more females in the classroom, in contrast to cognitive gains to boys, would suggest that changes in perceived competitiveness of the environment from more girls might plausibly be a significant contributory mechanism driving our results.

***Teacher Behaviour***

The fourth suggested route found in literature is via changes in teacher behaviour in response to peer gender composition. These include teachers adopting more interactive teaching styles like introducing more discussions amongst students and between themselves and the students (Lavy & Schlosser, 2011), allocating more time to tasks related to teaching and grading, and being generally more patient with and responsible for their students (Gong et al., 2019). Borbely et al. (2022) observe that if increasing the presence of females in the class causes the teacher to shift her attention toward the girls or causes a reduction in time and attention of the teacher spent on the boys in favour of the girls, both would cause girls' outcomes to improve based on Gong et al. (2019) and Lavy and Sand (2018).

Restricted by data, we only have limited means of testing out this mechanism. As in the case of students, we also have the math teachers' time allocation across positive and negative activities in the class available from the School Survey 1 of the YL. We contrast two variables capturing the proportion of positive activities by total, and negative activities by total activities engaged into by the teacher during the period of observation, and we regress each on the proportion of girls in the classroom and school fixed effects. We do not find any significant results and therefore, find that we cannot conclusively say that teachers' behaviour is the necessary mechanism of peer gender effect at play here. Having said that, we also cannot conclusively rule it out as a plausible mechanism either.

Our findings that girls improve in abilities across the subjects, which have different teachers, and boys benefit too, even though female-focus of the teacher does not predict that as per studies cited above, suggest that it is at least not the sole mechanism driving the result.

***Differences in approaches to learning***

Literature reports that there is a pervasive view that female students, especially, learn better in the company of other girls, and often parents' choice of single sex schools is also driven by the same rationale (Sohn, 2016). Further, evidences from Roberson Hayes et al. (2011) & Doris et al. (2013) suggest that "boys tend to both demand and receive more attention from teachers in classrooms", especially in male-dominated subjects like math and science, and therefore, an increase in the proportion of females in the classroom would effectively act to counterweigh that male-focus. We propose that this might be another reason why girls experience an improvement in mathematics as an outcome of increase in females in the class.

***Effects of gender stereotyping***

Amongst various perceived pros of single-sex education include arguments that students in single-sex institutions are protected from stereotyping and, therefore, stereotype threats, and therefore are free to pursue more non-stereotypical curricula and courses (Mael et al., 2004; Sohn, 2016). The above argument has also been extended in favour of single-sex education to imply that sexist attitudes harboured by peers effectively interfere with the learning outcomes of girls, as females are exposed to more stereotype susceptibility (Sohn, 2016; Sullivan et al., 2010; Scheeweis & Zweimuller, 2012). Shurchkov (2012) shows that when gender stereotypes are made less prominent, females are as willing to compete as males (Eisenkopf et al., 2015). Combining this with the common stereotype stating that math is a "male" subject, would suggest that increasing the proportion of females can effectively allay stereotype threats thereby improving learning in females, especially in math.

On the other hand, there is evidence to suggest that limited exposure to the opposite sex and sex segregation can in fact exacerbate gender stereotyping, legitimizing sexism (Halpern et al., 2011). Thus, increasing proportion of females in the classroom can plausibly have either positive or negative impact on the learning outcomes of both males and females by either exacerbating or mitigating the effects of sexist stereotypes.

In our particular context – a developing country where gender norms are especially strong (Borbely et al., 2022) – we think that the positive effects of the lessening of gender stereotype threat from having more females in the class, by making the stereotype of math being a "male" subject less prominent, might be an important factor causing an improvement in the cognitive abilities of the females (Eisenkopf et al., 2015). This stereotype threat erosion would logically happen in math and not the languages, and therefore that we record an achievement premium in math for females supports our reasoning of lessening of stereotype threat being a causal mechanism for the peer gender effect on the females. We do not believe the stereotypes are dominant in driving the male English ability improvement outcomes in our study.

## 6. Conclusion

In this study, we estimate the impact of the gender composition of primary school peers of Indian students on their Math, English and Telugu cognitive abilities using Young Lives survey data from Rounds 3 and School Survey 1. Our identifying assumption is that, conditional on various covariates, measures of lagged ability and school fixed effects, peer gender composition is de facto idiosyncratic across classes, and uncorrelated with the error term.

We find that a ten percentage point increase in the proportion of girls in the classroom improves math abilities of girls by about ten percent of a standard deviation, while a corresponding increase in female peers improves the English abilities of boys by 13.8 percent of a standard deviation. When we limit our analysis to coeducational schools, we additionally find that girls' English abilities also improve by 11.5 percent of a standard deviation while their math abilities also improve by nearly the same magnitude. Our results are robust to various tests, and randomized inference p-values as well as bound analysis following Oster (2013; 2019) confirm the same.

We argue that our results indicate that peer gender effects are mediated through lowering in disruptive classroom environment, prevalence of gendered social norms pertaining to inter-gender interactions, establishment of more female friendly competitive environments, perhaps changes in teacher behaviour, and most certainly because of reduction in stereotype salience which most benefits girls in a context like India which has demonstrably rigid gender norms and stereotypes against females.

Additionally, upon conducting heterogeneity analyses, we find that our results are driven by students who go to public schools as opposed to the more fee-charging private unaided schools, students who study in schools that do not offer instruction in English, and students who study in schools that have very limited resources and cannot, or do not, even provide children with gender segregated toilets. Thus, our results indicate that the greatest benefit from improving peer gender composition such that there is a higher proportion of girls in their peerage accrues to the least privileged students and those who depend largely on the government for the provision of primary education.

Lastly, we would like to preface all our findings by pointing out they were estimated using a pro-poor sample from a single state in India. Nevertheless, to the extent that the state of primary education in India is very similar across states, especially for government schools, our findings are generalizable to all of India, at least, and perhaps, if not in magnitude but in effect also apply to other developing countries with similar socio-economic-institutional contexts. At a minimum, our results from this study highlight that the effect of peer gender composition is positive, and heterogeneous for boys and girls across different domains of cognitive abilities.

References –

Alan, S., & Ertac, S. (2018). Fostering patience in the classroom: Results from randomized educational intervention. Journal of Political Economy, 126(5), 1865-1911.

Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality Psychology and Economics. In *Handbook of the Economics of Education* (Vol. 4, pp. 1–181). Elsevier. https://doi.org/10.1016/B978-0-444-53444-6.00001-8

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, *113*(1), 151–184.

Anelli, M., & Peri, G. (2019). The Effects of High School Peers' Gender on College Major, College Performance and Income. *The Economic Journal*, *129*(618), 553–602.

Angrist, J. D., & Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. *American Economic Review*, *94*(5), 1613-1634.

Antecol, H., Eren, O., & Ozbeklik, S. (2016). Peer effects in disadvantaged primary schools: Evidence from a randomized experiment. *Journal of Human Resources*, *51*(1), 95-132.

ASER (2023). Beyond Basics, Annual Status of Education Report 2023. https://asercentre.org/aser-2023-beyond-basics/

ASER. (2012). *Annual Status of Education Report (Rural) 2011*. https://img.asercentre.org/docs/Publications/ASER%20Reports/ASER_2011/aser_2011_report_8.2.12.pdf

Bengtsson, C., Persson, M., & Willenhag, P. (2005). Gender and overconfidence. *Economics letters*, *86*(2), 199-203.

Bertrand, M., & Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American economic journal: applied economics*, *5*(1), 32-64.

Betts, J. R., & Zau, A. (2004). Peer groups and academic achievement: Panel evidence from administrative data. *Unpublished manuscript*.

Billger, S. M. (2009). On reconstructing school segregation: The efficacy and equity of single-sex schooling. *Economics of education review*, *28*(3), 393-402.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2013). Under pressure? The effect of peers on outcomes of young adults. *Journal of Labor Economics*, *31*(1), 119-153.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2013). Under Pressure? The Effect of Peers on Outcomes of Young Adults. *Journal of Labor Economics*, *31*(1), 119–153.

Booth, A. & Nolen, P. (2012b), "Choosing to Compete: How Different Are Girls and Boys?", Journal of Economic Behavior and Organization, vol. 81, no. 2, pp. 542-555.

Booth, A. L. and Nolen, P. (2012a), Gender differences in risk behaviour: does nurture matter?. The Economic Journal, 122: F56F78

Booth, A., Cardona-Sosa, L., & Nolen, P., (2013). Do single-sex classes affect exam scores? An experiment in a coeducational university. Australian National University discussion paper, DP. No. 679.

Booth, Alison L, Cardona Sosa, Lina & Nolen, Patrick, 2014. "Do Single-Sex Classes Affect Achievement? An Experiment in a Coeducational University", CEPR Discussion Papers 10221, C.E.P.R. Discussion Papers

Boozer, M.A., Cacciola, S.E., 2001. Inside the black box of project star: Estimation of peer effects. Economic Growth Center Discussion Paper No. 832.

Borbely, D., Norris, J., & Romiti, A. (2022). *Peer gender and schooling: Evidence from Ethiopia*.

Brenøe, A. A., & Zölitz, U. (2020). Exposure to More Female Peers Widens the Gender Gap in STEM Participation. *Journal of Labor Economics*, *38*(4), 1009–1054.

Briole, S., 2021. Are girls always good for boys? Short and long term effects of school peers' gender. Econ. Educ. Rev. 84, 102–150.

Bryk, A., Lee, V., & Holland, P. (1993). Catholic schools and the common good. Cambridge, MA: Harvard University Press.

Burke, M.A., Sass, T.R., 2008. Classroom peer effects and student achievement. Calder Urban Institute Working Paper 18.

Canaan, S., & Mouganie, P. (2023). The impact of advisor gender on female students' STEM enrollment and persistence. *Journal of Human Resources*, *58*(2), 593-632.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, *134*(3), 1163-1224.

Carrell, S. E., & Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, *2*(1), 211-228.

Carrell, S. E., Fullerton, R. L., & West, J. E. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, *27*(3), 439-464.

Chandra, S. (2017). Primary education and minority rights. *Economic and Political Weekly, 52(46)*, 28-29.

Cho, R. M. (2012). Are there peer effects associated with having English language learner (ELL) classmates? Evidence from the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K). *Economics of Education Review*, *31*(5), 629-643.

Coleman, J. S. (1968). Equality of educational opportunity. *Integrated education*, *6*(5), 19-28.

Das, J., & Zajonc, T. (2010). India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics*, *92*(2), 175–187. https://doi.org/10.1016/j.jdeveco.2009.03.004

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human resources*, *42*(3), 528-554.

Dewan, P., Ray, T., Roy Chaudhuri, A., & Tater, K. (2024). Gender peer effects in high schools: Evidence from India. *Journal of Economic Behavior & Organization*, *220*, 470–494. https://doi.org/10.1016/j.jebo.2024.02.002

Doris, A., O'Neill, D., & Sweetman, O. (2013). Gender, single-sex schooling and maths achievement. *Economics of Education Review*, *35*, 104–119. https://doi.org/10.1016/j.econedurev.2013.04.001

Dunn, L. M., & Dunn, L. M. (1997). *Examiner's manual for the PPVT-III peabody picture vocabulary test: Form IIIA and Form IIIB*. AGS.

Eisenkopf, G., 2010. Peer effects, motivation, and learning. Econ. Educ. Rev. 29 (3), 364–374.

Eisenkopf, G., Hessami, Z., Fischbacher, U., & Ursprung, H. W. (2015). Academic performance and single-sex schooling: Evidence from a natural experiment in Switzerland. *Journal of Economic Behavior & Organization*, *115*, 123–143. https://doi.org/10.1016/j.jebo.2014.08.004

Epple, D., & Romano, R. E. (2011). Peer Effects in Education. In *Handbook of Social Economics* (Vol. 1, pp. 1053–1163). Elsevier. https://doi.org/10.1016/B978-0-444-53707-2.00003-7

Fryer, R.G., Levitt, S.D., 2010. An empirical analysis of the gender gap in mathematics. Am. Econ. J.: Appl. Econ. 2 (2), 210–240.

Gazta, P., & Jadhav, J. (2023). Does Access to Educational Institutions Signify Gaining Quality Education?. *Economic & Political Weekly*, *58*(30), 21.

Getik, D., Meier, A.N., 2022. Peer gender and mental health. J. Econ. Behav. Organ. 197, 643–659.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. The quarterly journal of economics, 118(3), 1049-1074.

Gong, J., Lu, Y., & Song, H. (2016). Gender Peer Effects on Students' Academic and Noncognitive Outcomes: Evidence and Mechanisms. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2844628

Guarino, C. M., Reckase, M. D., & Wooldridge, J. M. (2011). Evaluating Value-Added Methods of Estimating of Teacher Performance. *Society for Research on Educational Effectiveness*.

Hanushek, E.A., 1979. Conceptual and empirical issues in the estimation of educational production functions. J. Hum. Resour. 14 (3), 351–388.

Hanushek, E.A., Kain, J.F., Markman, J.M., Rivkin, S.G., 2003. Does peer ability affect student achievement? J. Appl. Econom. 18 (5), 527–544.

Hanushek, E.A., Kain, J.F., Rivkin, S.G., 2009. New evidence about Brown v. Board of Education: the complex effects of school racial composition on achievement. J. Labor Econ. 27 (3), 349–383.

Heß, Simon, "Randomization inference with Stata: A guide and software" *Stata Journal* 17(3) pp.

Hill, A.J., 2015. The girl next door: the effect of opposite gender friends on high school achievement. Am. Econ. J. Appl. Econ. 7 (3), 147–177.

Hoxby, C.M., 2000a. The effects of class size on student achievement: New evidence from population variation. Q. J. Econ. 15, 1239–1285.

Hoxby, C.M., 2000b. Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper No. 7867.

Hoxby, C.M., Weingarth, G., 2005. Taking race out of the equation: School reassignment and the structure of peer effects. Mimeograph.

Hu, F. (2015). Do girl peers improve your academic performance? *Economics Letters*, *137*, 54–58. https://doi.org/10.1016/j.econlet.2015.10.025

Imberman, S., Kugler, A. D., & Sacerdote, B. (2009). *Katrina's children: evidence on the structure of peer effects from hurricane evacuees* (No. w15291). National Bureau of Economic Research.

Imberman, S.A., Kugler, A.D., Sacerdote, B.I., 2012. Katrina's children: evidence on the structure of peer effects from hurricane evacuees. Am. Econ. Rev. 102 (5), 2048–2082.

Jackson, C. K. (2012). Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in Trinidad and Tobago. *Journal of Public Economics*, *96*(1–2), 173–187. https://doi.org/10.1016/j.jpubeco.2011.09.002

Jackson, C. K. (2021). Can Introducing Single-Sex Education into Low-Performing Schools Improve Academics, Arrests, and Teen Motherhood? *Journal of Human Resources*, *56*(1), 1–39. https://doi.org/10.3368/jhr.56.1.0618-9558R2

Jain, T., & Kapoor, M. (2015). The Impact of Study Groups and Roommates on Academic Performance. *Review of Economics and Statistics*, *97*(1), 44–54. https://doi.org/10.1162/REST_a_00454

James, Z., & Woodhead, M. (2014). Choosing and changing schools in India's private and government sectors: Young Lives evidence from Andhra Pradesh. *Oxford Review of Education*, *40*(1), 73-90.

Joensen, J. S., & Nielsen, H. S. (2016). Mathematics and gender: Heterogeneity in causes and consequences. *The Economic Journal*, *126*(593), 1129-1163.

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the national Academy of Sciences*, *103*(27), 10155-10162.

Lavy, V., & Sand, E. (2018). On the origins of gender gaps in human capital: Short-and long-term consequences of teachers' biases. *Journal of Public Economics*, *167*, 263-279.

Lavy, V., & Schlosser, A. (2007). Mechanisms and impacts of gender peer effects at school.

Lavy, V., & Schlosser, A. (2011). Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics*, *3*(2), 1–33. https://doi.org/10.1257/app.3.2.1

Lazear, E. P. (2001). Educational production. *The Quarterly Journal of Economics*, *116*(3), 777-803.

Lee, V. E., & Lockheed, M. E. (1990). The effects of single-sex schooling on achievement and attitudes in Nigeria. *Comparative Education Review*, *34*(2), 209-231.

Leon, J. (2020). Equating Cognitive Scores across Rounds and Cohorts for Young Lives in Ethiopia, India, Peru and Vietnam. *Young Lives.*

LoGerfo, L., Nichols, A., & Chaplin, D. (2006). Gender gaps in math and reading gains during elementary and high school by race and ethnicity. Urban Institute.

Lu, F., & Anderson, M. L. (2015). Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups. *Journal of Labor Economics*, *33*(1), 91–122. https://doi.org/10.1086/677392

Mael, F., Alonso, A., Gibson, D., Rogers, K., Smith, M., 2005. Single-sex versus coeducational schooling: a systematic review. American Institutes for Research. Technical report.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, *60*(3), 531-542.

Moffitt, R., 2001. Policy Interventions, Low-Level Equilibria, and Social Interactions. In: Blume, Durlauf (Eds.), Social Dynamics. Brookings Institution Press, Washington, DC, pp. 45–82.

Morrow, V., & Wilson, E. (2014). Parents' perspectives on quality of schooling in Andhra Pradesh, India. *Young Lives Working Paper. Young Lives, Department of International Development, University of Oxford*.

Nicole Schneeweis, Schneeweis, N., Martina Zweimüller, & Zweimüller, M. (2012). Girls, girls, girls: Gender composition and female school choice. *Economics of Education Review*, *31*(4), 482–500. https://doi.org/10.1016/j.econedurev.2011.11.002

Niederle, M. and Vesterlund, L. (2011), "Gender and Competition", Annual Review of Economics, vol. 3, no. 1, pp. 601-630.

Niederle, Muriel, Vesterlund, Lise (2007). “Do women shy away from competition? Do men compete too much?” Quarterly Journal of Economics 122 (3), 1067-1101.

Oosterbeek, H., Van Ewijk, R., 2014. Gender peer effects in university: evidence from a randomized experiment. Econ. Educ. Rev. 38, 51–63.

Oster, E. (2013). *Unobservable Selection and Coefficient Stability: Theory and Validation* (w19054; p. w19054). National Bureau of Economic Research. https://doi.org/10.3386/w19054

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204. https://doi.org/10.1080/07350015.2016.1227711

Park, H., Behrman, J. R., & Choi, J. (2013). Causal effects of single-sex schools on college entrance exams and college attendance: Random assignment in Seoul high schools. *Demography*, *50*, 447-469.

Rakshit, S., & Sahoo, S. (2023). Biased teachers and gender gap in learning outcomes: Evidence from India. *Journal of Development Economics*, *161*, 103041. https://doi.org/10.1016/j.jdeveco.2022.103041

Reddy, A. A., & Bantilan, M. (2012). Regional disparities in Andhra Pradesh, India. Local Economy the Journal of the Local Economy Policy Unit, 28(1), 123–135. https://doi.org/10.1177/0269094212463791.

Riordan, C. (1985). Public and catholic schooling: The effects of gender context policy. American Journal of Education, 93, 518–540.

Rivkin, S.G., 2001. Tiebout sorting, aggregation, and the estimation of peer group effects. Econ. Educ. Rev. 20, 201–209.

Roberson Hayes, A., Pahlke, E., & Bigler, R. (2011). The efficacy of single-sex education: Testing for selection and peer quality effects. Sex Roles: A Journal of Research, 65, 693–703.

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, *48*(2), 268-302.

Sacerdote, B. (2011). Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In *Handbook of the Economics of Education* (Vol. 3, pp. 249–277). Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00004-1

Sacerdote, B., 2001. Peer effects with random assignment: results for dartmouth roommates. Q. J. Econ. 116 (2), 681–704.

Schneeweis, N., & Zweimüller, M. (2012). Girls, girls, girls: Gender composition and female school choice. *Economics of Education review*, *31*(4), 482-500.

Shue, K. (2013). Executive networks and firm policies: Evidence from the random assignment of MBA peers. *The Review of Financial Studies*, *26*(6), 1401-1442.

Shurchkov, O. (2012). Under pressure: gender differences in output quality and quantity under competition and time constraints. Journal of the European Economic Association, 10(5), 1189-1213.

Singh, A. (2015). Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, *113*, 16–32. https://doi.org/10.1016/j.jdeveco.2014.10.004

Sohn, H. (2016). Mean and distributional impact of single-sex high schools on students' cognitive achievement, major choice, and test-taking behavior: Evidence from a random assignment policy in Seoul, Korea. *Economics of Education Review*, *52*, 155–175. https://doi.org/10.1016/j.econedurev.2016.02.007

Stehlé, J., Charbonnier, F., Picard, T., Cattuto, C., & Barrat, A. (2013). Gender homophily from spatial behavior in a primary school: A sociometric study. *Social Networks*, *35*(4), 604-613.

Stevenson, M. (2014). Breaking bad: social influence and the path to criminality in Juvenile Jails. *University of California, Berkeley, Department of Agricultural and Ressource Economics*.

Sullivan, C. J., Childs, K. K., & O'Connell, D. (2010). Adolescent risk behavior subgroups: An empirical assessment. Journal of youth and adolescence, 39, 541-562.

Vigdor, J.L., 2006. Peer effects in neighborhoods and housing. In: Dodge, K.A., Dishion, T.J., Lansford, J.E. (Eds.), Deviant Peer Influences in Programs for Youth: Problems and Solutions. Guilford Press.

Vigdor, J.L., Nechyba, T.S., 2007. Peer effects in North Carolina public schools. In: Woessmann, L., Peterson, P.E. (Eds.), Schools and the Equal Opportunity Problem. The MIT Press, pp. 73–101.

Winters, M. A., & Cowen, J. M. (2013). Who would stay, who would be dismissed? An empirical consideration of value-added teacher retention policies. *Educational Researcher*, *42*(6), 330-337.

World Bank, 2012. World development report.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and statistics*, *85*(1), 9-23.

## Tables

Table 1. Summary Statistics: Female vs male students

| | (1)<br>Female Students | (2)<br>Male Students | (3)<br>Difference |
|---|---|---|---|
| Panel A: Student Characteristics | | | |
| Age (in months) | 107.3 | 107.28 | -0.05 |
| *Child's Religion* | | | |
| Christian | 0.05 | 0.03 | -1.11 |
| Muslim | 0.05 | 0.06 | 1.03 |
| Hindu | 0.9 | 0.9 | 0.09 |
| Other (including tribal) | 0.01 | 0.01 | -0.55 |
| *Child's Mother Tongue* | | | |
| Telugu | 0.84 | 0.83 | -0.51 |
| Other than Telugu | 0.16 | 0.17 | 0.51 |
| *Child's Caste* | | | |
| SC (Scheduled Caste) | 0.2 | 0.19 | -0.4 |
| ST (Scheduled Tribe) | 0.11 | 0.13 | 0.69 |
| BC (Backward Caste) | 0.52 | 0.5 | -0.54 |
| OC (General or Other Caste) | 0.17 | 0.18 | 0.54 |
| *Child is firstborn or not* | | | |
| No | 0.65 | 0.68 | 0.78 |
| Yes | 0.35 | 0.32 | -0.78 |
| Panel B: Household Characteristics | | | |
| *Region of Andhra* | | | |
| Coastal Andhra | 0.35 | 0.3 | -1.58 |
| Rayalseema | 0.33 | 0.31 | -0.59 |
| Telangana | 0.32 | 0.38 | 2.13 |
| *Located in* | | | |
| Urban | 0.21 | 0.18 | -1.16 |
| Rural | 0.79 | 0.82 | 1.16 |
| *Mother's education* | | | |
| No education | 0.45 | 0.5 | 1.67 |
| Primary School (Class I-V) | 0.25 | 0.25 | -0.1 |
| Upper Primary School (Class VI-VII) | 0.11 | 0.09 | -1.09 |
| High School (Class VIII-X) | 0.15 | 0.13 | -0.91 |
| Junior College (Class XI-XII) | 0.03 | 0.01 | -2.33 |
| Higher education (E.g. University, Diploma) | 0.02 | 0.03 | 1.02 |
| *Father's education* | | | |
| No education | 0.27 | 0.31 | 1.35 |
| Primary School (Class I-V) | 0.25 | 0.24 | -0.41 |
| Upper Primary School (Class VI-VII) | 0.1 | 0.12 | 0.58 |
| High School (Class VIII-X) | 0.21 | 0.17 | -1.38 |
| Junior College (Class XI-XII) | 0.05 | 0.06 | 1.19 |
| Higher education (E.g. University, Diploma) | 0.06 | 0.06 | 0.26 |
| Number of Siblings | 1.64 | 1.62 | -0.32 |
| *Household Wealth Quartiles* | | | |
| Lowest | 0.3 | 0.23 | -2.18 |
| Second Lowest | 0.25 | 0.28 | 0.83 |

| | | | |
|---|---|---|---|
| Middle | 0.26 | 0.28 | 0.64 |
| Highest | 0.19 | 0.21 | 0.76 |
| Time taken to reach school | 10.58 | 11.59 | 1.59 |
| Panel C: School Characteristics | | | |
| *Type of School* | | | |
| Government School | 0.69 | 0.58 | -3.46*** |
| Private Aided School | 0.02 | 0.02 | -0.53 |
| Private Unaided School | 0.29 | 0.41 | 3.66*** |
| *Medium of Instruction* | | | |
| Does not provide instruction in English | 0.72 | 0.62 | -2.98** |
| Provides instruction in English | 0.28 | 0.38 | 2.98** |
| *Gender Segregated Toilets* | | | |
| No | 0.59 | 0.6 | 0.2 |
| Yes | 0.41 | 0.4 | -0.2** |
| Panel D: Class and Teacher Characteristics | | | |
| Student Teacher Ratio in the class | 24.78 | 26.35 | 2.07 |
| *Classroom Gender Norm* | | | |
| Girls and boys not required to sit separately | 0.34 | 0.3 | -1.3 |
| Girls and boys required to sit separately | 0.66 | 0.7 | 1.3 |
| *Math Teacher's Gender* | | | |
| Male | 0.6 | 0.57 | -0.93 |
| Female | 0.4 | 0.43 | 0.93 |
| *Math Teacher's Caste* | | | |
| SC (Scheduled Caste) | 0.16 | 0.21 | 1.83 |
| ST (Scheduled Tribe) | 0.08 | 0.07 | -0.35 |
| BC (Backward Caste) | 0.48 | 0.42 | -1.77 |
| OC (General or Other Caste) | 0.28 | 0.3 | 0.58 |
| *Math Teacher's Highest Qualification* | | | |
| Matriculation Passed (10th) | 0.02 | 0.04 | 2.25 |
| Higher Secondary Passed (12th) | 0.18 | 0.2 | 0.8 |
| Graduation (Bachelor) | 0.57 | 0.55 | -0.83 |
| Masters (Other Post Graduation) | 0.23 | 0.21 | -0.72 |
| *Math Teacher's Employment Contract* | | | |
| Permanent (or regular) | 0.58 | 0.46 | -3.58*** |
| Temporary | 0.42 | 0.54 | 3.58*** |
| Math Teacher's Gender Attitude Index | -0.03 | 0.1 | 1.9 |
| Math Teacher's Age | 32.13 | 31.15 | -1.66 |
| Panel E: Average Classroom Peer Ability - Endline | | | |
| Average Math Score of the Class | 63.39 | 66.19 | 2.69** |
| Average Telugu Score of the Class | 65.42 | 68.41 | 2.84*** |
| Average English Score of the Class | 61.06 | 65.43 | 4.00*** |
| N | 444 | 511 | 955 |

*Notes: Column 3 shows the difference in characteristics between female and male students with t-test results. (3) = (2)-(1). ***p<0.01, **p<0.05*

Table 2. Balance Tests – Coefficients on the share of female peers in the class

| | (1) Full | (2) Full | (3) Male | (4) Male | (5) Female | (6) Female |
|---|---|---|---|---|---|---|
| PANEL A: Student Characteristics | | | | | | |
| *Baseline Ability Scores (Round 3 scores)* | | | | | | |
| Math | -0.01 | 0.01 | -0.01 | 0.00 | -0.00 | 0.02 |
| | (0.007) | (0.007) | (0.009) | (0.010) | (0.009) | (0.011) |
| PPVT | 0.00 | 0.02*** | 0.01 | 0.01 | 0.00 | 0.02** |
| | (0.007) | (0.007) | (0.008) | (0.009) | (0.011) | (0.011) |
| Agency | -0.01** | 0.00 | -0.01* | 0.01 | -0.01 | 0.01 |
| | (0.005) | (0.005) | (0.007) | (0.008) | (0.008) | (0.009) |
| Pride and Self-esteem | -0.02*** | -0.00 | -0.02** | -0.00 | -0.01 | 0.01 |
| | (0.006) | (0.005) | (0.008) | (0.008) | (0.008) | (0.009) |
| Child's age | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.002) | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) |
| *Region (Ref gp.: Coastal Andhra)* | | | | | | |
| Rayalseema | -0.04 | -0.13*** | -0.04 | 0.10*** | -0.03 | -0.17* |
| | (0.022) | (0.026) | (0.024) | (0.025) | (0.027) | (0.101) |
| Telangana | -0.02 | 0.08 | -0.03 | 0.06 | -0.01 | -0.08 |
| | (0.020) | (0.065) | (0.021) | (0.058) | (0.030) | (0.030) |
| Urban | 0.00 | 0.07* | 0.01 | -0.01 | -0.02 | 0.17* |
| | (0.015) | (0.039) | (0.017) | (0.057) | (0.020) | (0.101) |
| *Child's Religion (Ref gp.: Other)* | | | | | | |
| Christian | 0.04 | 0.00 | -0.04 | 0.03 | 0.10* | -0.02 |
| | (0.044) | (0.036) | (0.072) | (0.046) | (0.050) | (0.034) |
| Muslim | 0.04 | -0.01 | -0.03 | 0.04 | 0.11** | -0.04 |
| | (0.044) | (0.035) | (0.071) | (0.049) | (0.052) | (0.035) |
| Hindu | -0.01 | -0.01 | -0.06 | 0.05 | 0.04 | -0.04 |
| | (0.033) | (0.026) | (0.064) | (0.030) | (0.027) | (0.023) |
| *Child's Caste (Ref gp.: General/Other Castes)* | | | | | | |
| SC | 0.02 | -0.00 | 0.03 | -0.01 | 0.00 | 0.02 |
| | (0.021) | (0.015) | (0.026) | (0.021) | (0.027) | (0.021) |
| ST | 0.05 | -0.02 | 0.04 | -0.07 | 0.06 | 0.02 |
| | (0.031) | (0.024) | (0.031) | (0.047) | (0.055) | (0.031) |
| BC | 0.00 | -0.01 | 0.00 | 0.01 | -0.00 | 0.01 |
| | (0.017) | (0.013) | (0.021) | (0.021) | (0.023) | (0.019) |
| Mother Tongue – Telugu | -0.01 | 0.02 | -0.01 | 0.06 | 0.00 | 0.01 |
| | (0.020) | (0.024) | (0.021) | (0.046) | (0.027) | (0.022) |
| PANEL B: Household Characteristics | | | | | | |
| *Mother's education (Ref gp.: No education)* | | | | | | |
| Primary School (Class I-V) | 0.02 | 0.03** | 0.00 | 0.03* | 0.03 | 0.03 |
| | (0.015) | (0.014) | (0.019) | (0.019) | (0.023) | (0.022) |
| Upper Primary School (Class VI-VII) | 0.01 | -0.01 | -0.02 | 0.02 | 0.04 | -0.00 |
| | (0.019) | (0.014) | (0.022) | (0.023) | (0.028) | (0.021) |
| High School (Cass VIII-X) | -0.04** | 0.01 | -0.05** | 0.03 | -0.03 | 0.01 |
| | (0.017) | (0.016) | (0.020) | (0.020) | (0.024) | (0.028) |
| Junior College (Class XI-XII) | -0.04* | 0.05* | -0.01 | 0.11 | -0.06** | 0.05 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | (0.022) | (0.028) | (0.038) | (0.083) | (0.027) | (0.035) |
| Higher Education (e.g. University, Diploma) | -0.02 | 0.01 | -0.05 | 0.00 | 0.04 | 0.00 |
| | (0.038) | (0.022) | (0.035) | (0.024) | (0.075) | (0.041) |
| *Father's education (Ref gp.: No education)* | | | | | | |
| Primary School (Class I-V) | 0.01 | 0.00 | 0.01 | 0.02 | -0.00 | -0.01 |
| | (0.017) | (0.013) | (0.019) | (0.017) | (0.028) | (0.027) |
| Upper Primary School (Class VI-VII) | -0.01 | -0.00 | -0.02 | -0.03 | -0.01 | -0.00 |
| | (0.018) | (0.015) | (0.023) | (0.024) | (0.029) | (0.028) |
| High School (Class VIII-X) | -0.05*** | -0.01 | -0.07*** | -0.02 | -0.04 | 0.02 |
| | (0.017) | (0.014) | (0.021) | (0.019) | (0.026) | (0.024) |
| Junior College (Class XI-XII) | -0.02 | 0.04* | -0.02 | 0.05* | -0.03 | 0.01 |
| | (0.023) | (0.022) | (0.027) | (0.029) | (0.042) | (0.037) |
| Higher Education (e.g. University, Diploma) | -0.04* | 0.01 | -0.04 | 0.02 | -0.05 | 0.01 |
| | (0.022) | (0.018) | (0.030) | (0.025) | (0.029) | (0.029) |
| Not known | -0.01 | -0.00 | 0.00 | 0.02 | -0.03 | -0.02 |
| | (0.028) | (0.025) | (0.048) | (0.042) | (0.030) | (0.027) |
| Number of siblings | -0.00 | -0.01 | 0.00 | -0.01* | -0.00 | 0.00 |
| | (0.005) | (0.006) | (0.007) | (0.008) | (0.007) | (0.009) |
| Firstborn | -0.00 | -0.00 | -0.01 | -0.03* | 0.00 | -0.01 |
| | (0.013) | (0.011) | (0.016) | (0.015) | (0.019) | (0.018) |
| *Wealth Index Quartiles (Ref gp.: 4)* | | | | | | |
| Wealth quartile 1 | 0.05** | -0.02 | 0.04** | 0.01 | 0.04 | -0.04 |
| | (0.019) | (0.021) | (0.021) | (0.027) | (0.030) | (0.037) |
| Wealth quartile 2 | 0.03 | -0.01 | 0.03 | -0.01 | 0.02 | -0.00 |
| | (0.017) | (0.019) | (0.021) | (0.024) | (0.024) | (0.034) |
| Wealth quartile 3 | -0.01 | -0.02 | -0.01 | -0.00 | -0.00 | -0.01 |
| | (0.014) | (0.018) | (0.018) | (0.022) | (0.021) | (0.033) |
| Time taken to reach school | -0.00*** | 0.00 | -0.00*** | -0.00 | -0.00 | 0.00 |
| | (0.001) | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) |
| School Fixed Effects | No | Yes | No | Yes | No | Yes |

*Notes: Each cell represents a separate balance test that regresses the indicated variable on the main treatment variable – leave-one-out proportion of females in the classroom. Columns (1) and (2) are run on the full sample, while columns (3-6) are run on the male and female samples, separately, as indicated in the column title.*
*Robust standard errors are clustered at the school-class level, and are reported in the parentheses.*
*$^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$*

Table 3. Validity Check: Sample Attrition

| | (1) Full Sample | (2) Full Sample | (3) Male | (4) Male | (5) Female | (6) Female |
|---|---|---|---|---|---|---|
| Attrition | -0.05 | 0.11 | -0.14 | 0.15 | 0.04 | -0.01 |
| | (0.138) | (0.137) | (0.175) | (0.231) | (0.177) | (0.224) |
| Observations | 950 | 950 | 506 | 506 | 444 | 444 |
| R-squared | 0.000 | 0.555 | 0.002 | 0.621 | 0.000 | 0.641 |
| School FE | No | Yes | No | Yes | No | Yes |

*Notes: Sample attrition is regressed on the main treatment variable – leave-one-out proportion of females in the class. Attrition dummy takes value 1 for any of the covariates (listed in full in Table 4) missing value for an observation, and zero otherwise.*
*Robust standard errors are clustered at the school-class level, and are reported in the parentheses.*
*$^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$*

Table 4. Main Results – Effects of Peer Gender Composition on Cognitive Abilities by Gender

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **PANEL A: FEMALES** | | | | | | | | |
| **(a) Math** | **0.57**** | **0.56**** | **0.70**** | **0.72**** | **0.87**** | **0.87**** | **1.04**** | **1.04**** |
| | (0.223) | (0.226) | (0.323) | (0.299) | (0.361) | (0.361) | (0.417) | (0.417) |
| [Randomized Inference p] | [0.013] | [0.019] | [0.021] | [0.024] | [0.011] | [0.009] | [0.004] | [0.006] |
| Observations | 440 | 406 | 406 | 363 | 352 | 352 | 351 | 351 |
| $R^2$ | 0.33 | 0.381 | 0.67 | 0.673 | 0.677 | 0.677 | 0.687 | 0.687 |
| **(b) English** | -0.3 | -0.03 | **0.70*** | **0.73*** | **0.83** | **0.83** | | |
| | (0.282) | (0.268) | (0.396) | (0.420) | (0.511) | (0.511) | | |
| [Randomized Inference p] | [0.219] | [0.903] | [0.004] | [0.015] | [0.008] | [0.003] | | |
| Observations | 418 | 387 | 387 | 345 | 336 | 336 | | |
| $R^2$ | 0.258 | 0.405 | 0.773 | 0.754 | 0.756 | 0.756 | | |
| **(c) Telugu** | 0.60*** | 0.62*** | 0.74 | 0.71 | 0.73 | 0.73 | | |
| | (0.229) | (0.233) | (0.474) | (0.477) | (0.582) | (0.582) | | |
| [Randomized Inference p] | [0.014] | [0.014] | [0.013] | [0.043] | [0.039] | [0.031] | | |
| Observations | 436 | 403 | 403 | 360 | 349 | 349 | | |
| $R^2$ | 0.318 | 0.371 | 0.694 | 0.709 | 0.707 | 0.707 | | |
| **PANEL B: MALES** | | | | | | | | |
| **(a) Math** | 0.34 | 0.32 | 0.67* | 0.59 | 0.55 | 0.55 | **0.98**** | **0.98**** |
| | (0.283) | (0.284) | (0.394) | (0.397) | (0.400) | (0.399) | (0.482) | (0.482) |
| [Randomized Inference p] | [0.139] | [0.172] | [0.015] | [0.060] | [0.068] | [0.054] | [0.003] | [0.003] |
| Observations | 504 | 449 | 449 | 396 | 393 | 392 | 390 | 390 |
| $R^2$ | 0.364 | 0.461 | 0.74 | 0.741 | 0.742 | 0.742 | 0.752 | 0.752 |
| **(b) English** | **-0.64*** | -0.05 | **1.12**** | **1.18**** | **1.38**** | **1.38**** | | |
| | (0.341) | (0.316) | (0.481) | (0.522) | (0.535) | (0.534) | | |
| [Randomized Inference p] | [0.013] | [0.836] | [0.000] | [0.000] | [0.000] | [0.000] | | |
| Observations | 484 | 429 | 429 | 381 | 378 | 377 | | |
| $R^2$ | 0.275 | 0.427 | 0.787 | 0.792 | 0.794 | 0.793 | | |
| **(c) Telugu** | 0.54** | 0.41* | 0.42 | 0.36 | 0.53 | 0.53 | | |
| | (0.237) | (0.230) | (0.398) | (0.431) | (0.448) | (0.447) | | |
| [Randomized Inference p] | [0.015] | [0.064] | [0.138] | [0.263] | [0.099] | [0.079] | | |
| Observations | 498 | 443 | 443 | 390 | 387 | 386 | | |
| $R^2$ | 0.345 | 0.445 | 0.696 | 0.69 | 0.697 | 0.697 | | |
| Lagged child ability | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Student controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Household controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School FE | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer ability controls | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Class controls | No | No | No | No | Yes | Yes | Yes | Yes |
| School controls | No | No | No | No | No | Yes | Yes | No |

*Notes: Student controls include child's age, grade, a dummy each for region (Coastal Andhra, Telangana or Rayalseema), type of site (urban or rural), religion (Christian, Hindu, Muslim, Tribal), caste (Scheduled Caste, Scheduled Tribe, Other Backward Castes, Other Castes) and mother tongue (Telugu or not) of the child, and household controls include mother's education, dummy for firstborn, number of siblings and number of siblings squared, wealth index, and time to reach school. School controls include a dummy for type of school (public, private unaided and aided), instruction in English, separate toilets for girls, and class controls include dummy for opposite sex*

*students are not allowed to sit next to each other. Teacher controls include class teacher's gender, caste, qualifications, age, nature of employment (temporary or permanent) and an index of teacher's gender attitude. Peer ability controls include average marks in Math, English and Telugu for the class. Standardized IRT values of cognitive abilities are used as dependent variables.*
*Robust standard errors are clustered at the school-class level, and are reported in the parentheses.*
*Randomized inference p-values are reported in square brackets.*
****p<0.01, **p<0.05, *p<0.1*

Table 5. Robustness Check – Restricted Sample Analysis

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **PANEL A: FEMALES** | | | | | | | | |
| **(a) Math** | **0.88*** ** | **0.87** ** | **0.95** ** | **0.97** ** | **0.99** ** | **0.99** ** | **1.16** ** | **1.16** ** |
| | (0.327) | (0.345) | (0.449) | (0.417) | (0.428) | (0.428) | (0.499) | (0.499) |
| [Randomized Inference p] | [0.000] | [0.000] | [0.021] | [0.022] | [0.015] | [0.025] | [0.012] | [0.009] |
| Observations | 419 | 386 | 386 | 344 | 343 | 343 | 342 | 342 |
| $R^2$ | 0.339 | 0.39 | 0.676 | 0.679 | 0.68 | 0.68 | 0.69 | 0.69 |
| **(b) English** | -0.42 | -0.14 | **1.08** ** | **1.14** ** | **1.15** ** | **1.15** ** | | |
| | (0.414) | (0.420) | (0.507) | (0.553) | (0.556) | (0.556) | | |
| [Randomized Inference p] | [0.183] | [0.643] | [0.001] | [0.003] | [0.004] | [0.002] | | |
| Observations | 400 | 370 | 370 | 329 | 329 | 329 | | |
| $R^2$ | 0.259 | 0.41 | 0.777 | 0.76 | 0.76 | 0.76 | | |
| **(c) Telugu** | 0.4 | 0.38 | 0.75 | 0.67 | 0.69 | 0.69 | | |
| | (0.334) | (0.344) | (0.682) | (0.704) | (0.707) | (0.707) | | |
| [Randomized Inference p] | [0.217] | [0.264] | [0.083] | [0.127] | [0.140] | [0.132] | | |
| Observations | 415 | 383 | 383 | 341 | 340 | 340 | | |
| $R^2$ | 0.312 | 0.366 | 0.692 | 0.708 | 0.709 | 0.709 | | |
| **PANEL B: MALES** | | | | | | | | |
| **(a) Math** | 0.35 | 0.25 | 0.62 | 0.52 | 0.55 | 0.55 | 0.98** | 0.98** |
| | 0.307 | 0.308 | 0.409 | 0.395 | 0.398 | 0.397 | 0.48 | 0.48 |
| [Randomized Inference p] | [0.158] | [0.311] | [0.046] | [0.118] | [0.097] | [0.107] | [0.004] | [0.007] |
| Observations | 497 | 443 | 443 | 391 | 389 | 388 | 386 | 386 |
| $R^2$ | 0.363 | 0.459 | 0.734 | 0.736 | 0.737 | 0.737 | 0.746 | 0.746 |
| **(b) English** | **-0.54** | **-0.03** | **1.27*** ** | **1.36** ** | **1.38*** ** | **1.38*** ** | | |
| | 0.343 | 0.363 | 0.485 | 0.528 | 0.532 | 0.531 | | |
| [Randomized Inference p] | [0.054] | [0.918] | [0.000] | [0.000] | [0.000] | [0.000] | | |
| Observations | 478 | 424 | 424 | 376 | 374 | 373 | | |
| $R^2$ | 0.279 | 0.428 | 0.788 | 0.793 | 0.793 | 0.792 | | |
| **(c) Telugu** | 0.75*** | 0.60** | 0.58 | 0.47 | 0.53 | 0.53 | | |
| | 0.26 | 0.253 | 0.417 | 0.449 | 0.445 | 0.445 | | |
| [Randomized Inference p] | [0.002] | [0.007] | [0.072] | [0.175] | [0.128] | [0.114] | | |
| Observations | 491 | 437 | 437 | 385 | 383 | 382 | | |
| $R^2$ | 0.347 | 0.447 | 0.694 | 0.689 | 0.696 | 0.696 | | |
| Lagged child ability | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Student controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Household controls | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School FE | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer ability controls | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Class controls | No | No | No | No | Yes | Yes | Yes | Yes |
| School controls | No | No | No | No | No | Yes | Yes | No |

*Notes: These regressions are run on the sample of students in coeducational schools only. The full list of controls are listed in the notes of Table 4. Standardized IRT values of cognitive abilities are used as dependent variables.*
*Robust standard errors are clustered at the school-class level, and are reported in the parentheses.*
*Randomized inference p-values are reported in square brackets. ***p<0.01, **p<0.05, *p<0.1*

Table 6: Oster Bounds – Testing Potential Bias on Unobservables

| | $\beta(R^2)$ | | Estimated Bias when $R_{max} = 1$ | | | |
|---|---|---|---|---|---|---|
| | Uncontrolled | Controlled | $\beta$ *for* $\delta = 1$ | $\beta$ *for* $\delta = 0.5$ | $\beta$ *for* $\delta = 0.25$ | $\delta$ *for* $\beta = 0$ |
| Females | | | | | | |
| Math | 0.608 (0.009) | 1.037 (0.687) | 7.294 | 1.641 | 1.245 | 2.449 |
| English | -0.313 (0.002) | 0.694 (0.770) | 2.953 | 1.361 | 0.972 | -1.333 |
| Telugu | 0.480 (0.005) | 0.741 (0.722) | 2.622 | 1.017 | 0.843 | 2.407 |
| Males | | | | | | |
| Math | 0.238 (0.001) | 0.982 (0.752) | 12.811 | 2.173 | 1.368 | 24.329 |
| English | -0.515 (0.006) | 1.545 (0.813) | 7.853 | 3.266 | 2.255 | -1.435 |
| Telugu | 0.564 (0.009) | 0.822 (0.712) | 48.843 | 1.666 | 1.013 | 1.125 |

*Notes: $R_{max}$ is taken to be 1 for these calculations (Oster, 2019). Uncontrolled regressions are conducted by regressing standardized IRT values of indicated cognitive ability on leave-one-out proportion of girls in the classroom, and school fixed effects. Controlled regressions include all individual, household, school, class and teacher controls (refer Table 4 for the full list), and school fixed effects.*

*$\delta = 0.25$ produces $\beta$ estimates that are closest to observed $\beta$s indicating that selection on unobservables is most likely much less than selection on observables. Negative $\delta$s indicate that selection on unobservables would have to be opposite in direction to selection on observables for $\beta$ to be zero.*

Table 7. Heterogeneity and Sub-sample Analysis

| | Math | | English | | Telugu | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Type of School** | **Public** | **Private** | **Public** | **Private** | **Public** | **Private** |
| **Females** | **0.92**** | 0.85 | 0.97 | -0.42 | 0.1 | 2.18 |
| | (0.379) | (2.028) | (0.599) | (1.871) | (0.532) | (2.726) |
| Observations | 239 | 103 | 224 | 102 | 239 | 100 |
| R-squared | 0.66 | 0.894 | 0.672 | 0.937 | 0.7 | 0.953 |
| **Males** | 0.76 | 4.47* | **2.07***** | 0.73 | 0.64 | 5.03 |
| | (0.498) | (2.558) | (0.526) | (1.464) | (0.508) | (3.218) |
| Observations | 241 | 143 | 226 | 143 | 239 | 139 |
| R-squared | 0.775 | 0.832 | 0.795 | 0.848 | 0.734 | 0.842 |
| **Panel B: Medium of Instruction English** | **No** | **Yes** | **No** | **Yes** | **No** | **Yes** |
| **Females** | **0.92**** | 0.77 | 0.91 | -0.2 | 0.09 | 2.11 |
| | (0.387) | (2.684) | (0.607) | (1.973) | (0.530) | (2.977) |
| Observations | 250 | 101 | 235 | 100 | 250 | 98 |
| R-squared | 0.662 | 0.904 | 0.673 | 0.952 | 0.703 | 0.954 |
| **Males** | 0.83* | 2.54 | **1.62***** | 1.65 | 0.62 | 4.92 |
| | (0.483) | (3.898) | (0.571) | (1.552) | (0.493) | (5.607) |
| Observations | 259 | 131 | 244 | 131 | 257 | 127 |
| R-squared | 0.784 | 0.785 | 0.786 | 0.896 | 0.73 | 0.828 |
| **Panel C: Availability of Toilets for Girls** | **No** | **Yes** | **No** | **Yes** | **No** | **Yes** |
| **Females** | 0.55 | **2.35*** | **1.58**** | -0.3 | **0.81*** | -1.02 |
| | (0.425) | (1.332) | (0.649) | (0.744) | (0.452) | (2.033) |
| Observations | 208 | 143 | 193 | 142 | 208 | 140 |
| R-squared | 0.675 | 0.799 | 0.739 | 0.898 | 0.752 | 0.821 |
| **Males** | 1.05* | -0.91 | 0.58 | 1.99 | 0.75 | 6.49 |
| | (0.540) | (3.392) | (0.478) | (1.948) | (0.507) | (4.584) |
| Observations | 239 | 152 | 225 | 151 | 237 | 148 |
| R-squared | 0.782 | 0.807 | 0.809 | 0.92 | 0.705 | 0.824 |

*Notes: Each cell represents a separate regression of indicated standardized IRT cognitive ability scores on leave-one-out proportion of girls in the classroom for the indicated sub-sample, for males and females separately, based on equation (8) for math and equation (6) for English and Telugu. The full list of controls are listed in the notes of Table 4.*
*Robust standard errors are clustered at school-class level, and reported in the parentheses.*
****p<0.01, **p<0.05, *p<0.1*

Table 8. Balance Tests for Peer Ability, Class and Math Teacher Characteristics

| | (1) Male | (2) Male | (1) Female | (2) Female |
|---|---|---|---|---|
| PANEL A: Average Peer Ability | | | | |
| Class average of Math scores | -0.00** | 0.00*** | -0.00** | -0.01 |
| | (0.001) | (0.001) | (0.001) | -0.01 |
| Class average of English scores | -0.00*** | 0.00*** | -0.00*** | -0.02*** |
| | (0.001) | (0.001) | (0.001) | (0.000) |
| Class average of Telugu scores | -0.00** | 0.03*** | -0.00** | -0.01*** |
| | (0.001) | (0.006) | (0.001) | (0.000) |
| PANEL B: Class Characteristics | | | | |
| Student to Teacher Ratio in the Class | -0.00*** | -0.01** | -0.00** | 0.03 |
| | (0.001) | (0.003) | (0.001) | -0.02 |
| Classroom gender norms – Girls and boys sit separately | -0.03 | 0.03 | -0.06** | -0.04 |
| | (0.023) | (0.030) | (0.024) | (0.042) |
| PANEL C: Math Teacher Characteristics | | | | |
| Teacher's gender attitude Index | -0.02*** | -0.01 | -0.02** | 0.01 |
| | (0.008) | (0.013) | (0.011) | (0.018) |
| Teacher is Female | -0.03 | -0.04 | 0.01 | 0.00 |
| | (0.019) | (0.030) | (0.024) | (0.042) |
| *Teacher's Caste (Reference group – General)* | | | | |
| SC (Scheduled Caste) | -0.00 | -0.00 | 0.01 | 0.05 |
| | (0.030) | (0.046) | (0.032) | (0.068) |
| ST (Scheduled Tribe) | 0.12*** | 0.12 | 0.11* | 0.15 |
| | (0.038) | (0.122) | (0.065) | (0.100) |
| BC (Backward Caste) | 0.05** | -0.01 | 0.02 | -0.01 |
| | (0.020) | (0.050) | (0.029) | (0.066) |
| *Teacher's highest qualifications (Ref. gp – Class X)* | | | | |
| Teacher's age (in years) | 0.00** | 0.00 | 0.00** | 0.00 |
| | (0.001) | (0.002) | (0.001) | (0.002) |
| Higher Secondary Passed (12th) | 0.01 | -0.00 | 0.10** | **0.21**** |
| | (0.042) | (0.023) | (0.040) | (0.099) |
| Graduation (Bachelor) | -0.02 | -0.02 | 0.05* | **0.19**** |
| | (0.040) | (0.032) | (0.028) | (0.094) |
| Masters (Other Post Graduation) | -0.04 | -0.01 | 0.06* | **0.26**** |
| | (0.046) | (0.050) | (0.036) | (0.108) |
| Teacher is permanent | -0.07*** | -0.00 | -0.05** | -0.03 |
| | (0.019) | (0.030) | (0.023) | (0.053) |
| School FE | No | Yes | No | Yes |

*Notes: Each cell represents a separate regression of the indicated characteristic against the main treatment variable – leave-one-out proportion of girls in the classroom, for the male (columns 1 and 2), and female (columns 3 and 4) samples separately.*
*Robust standard errors are clustered at school-class level, and reported in the parentheses.*
****p<0.01, **p<0.05, *p<0.1*

## Figures

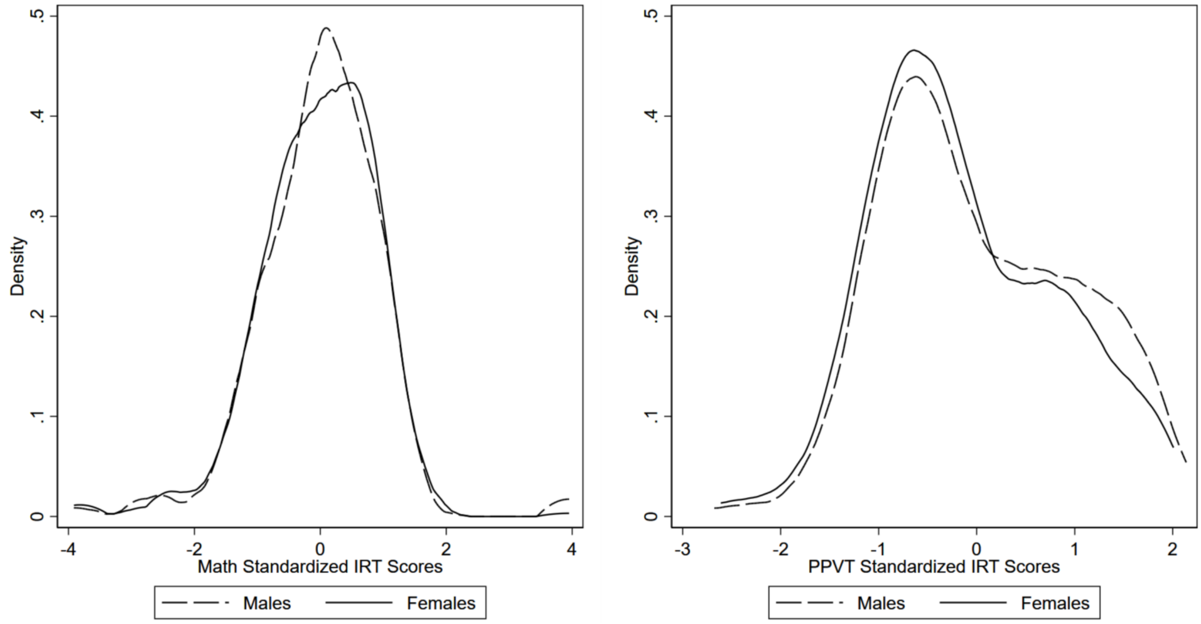


***Figure 1.*** Distribution of math and PPVT scores by gender of the student. Notes: Standardized IRT math and PPVT scores are plotted by gender. Math and PPVT are baseline scores taken from Round 3 survey of the Young Lives in 2009.

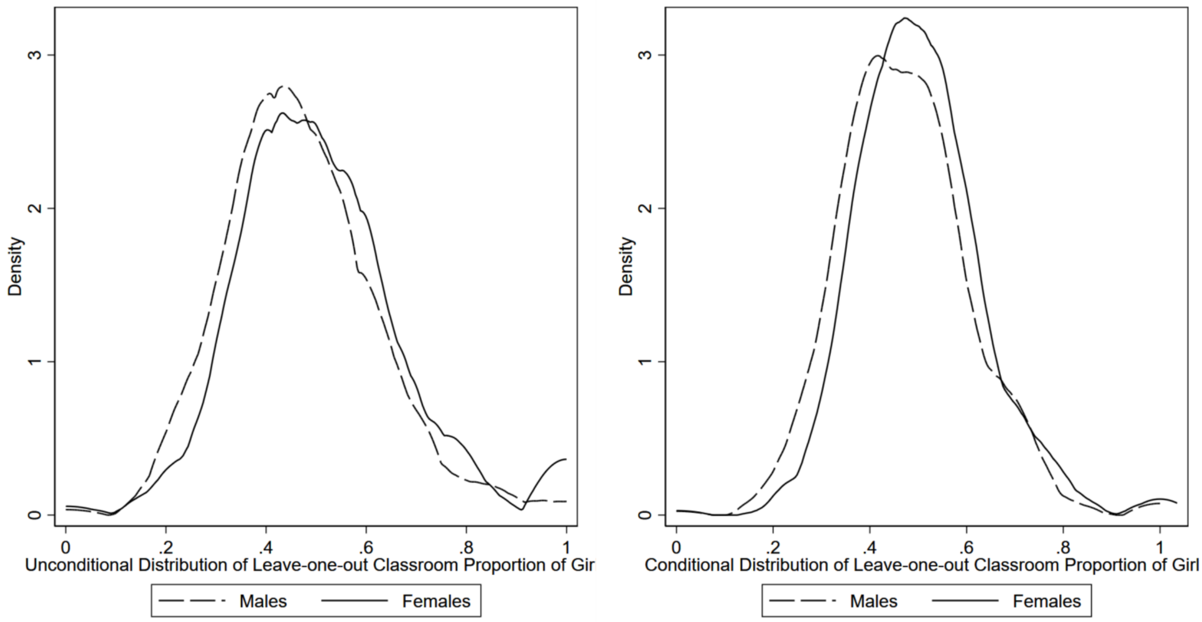


***Figure 2.*** Distribution of leave-one-out proportion of girls in the classroom – unconditional and conditional – by gender