Gender Peer Effects on Students' Academic and Noncognitive Outcomes: Evidence and Mechanisms

Jie Gong, Yi Lu and Hong Song

National University of Singapore, Singapore*

This Version: August 2016

Abstract

This paper estimates the gender peer effects on students' academic performance and noncognitive outcomes. The analysis uses a nationally representative survey of middle school students in China and focuses on schools that randomly assign students into classrooms. Our findings show that having a higher proportion of female peers in class improves students' test scores and noncognitive outcomes, such as mental stress, social acclimation and satisfaction, and disciplinary problems. There is evidence in support of improved classroom environment, teacher-student interactions, and teachers' level of job satisfaction when there are more female students in the class. The findings extend the study of gender peer effects from the traditional focus on academic achievements to noncognitive outcomes, and provide useful evidence on potential mechanisms.

Keyword: peer effects; gender; noncognitive outcomes

JEL Classification: I21, J16, Z13

*Gong (gong@nus.edu.sg): NUS Business School, Department of Strategy and Policy. 15 Kent Ridge Drive, Singapore 119245. Lu (ecsluyi@nus.edu.sg) and Song (songhong@u.nus.edu): Department of Economics, 1 Arts Singapore 117570.

1 Introduction

This paper investigates how the gender composition of the class affects students' academic and noncognitve outcomes. Researchers and policymakers have long believed that peer effects—e.g., gender, race, quality and social background—are among the most important determinants of student outcomes (e.g., Sacerdote 2001, Zimmerman 2003, Angrist and Lang 2004, Arcidiacono and Nicholson 2005, Ammermueller and Pischke 2009, Carrell, Fullerton and West 2009, Gould, Lavy and Daniele Paserman 2009).[1] Understanding the interaction of gender in the educational production function is particularly relevant for optimal grouping of students into schools and classrooms, and may shed light on the debate between single-sex and co-educational schools. Along this line, previous studies have highlighted the influence of the presence of girls on peers' academic outcomes. For example, Hoxby (2000) shows that a higher proportion of girls in the class raises students test scores in public elementary schools in Texas. Lavy and Schlosser (2011) extend the level of education to middle and high schools and further uncover the mechanisms—improved classroom environment and better inter-student and teacher-student relationships—behind the positive externalities. Black, Devereux and Salvanes (2013) examine the role of peer composition on students' long run outcomes, and find that having a higher proportion of girls in the class reduces boys' probability onto an academic high school track and their long-term completed years of schooling.

In contrast to the established effects on students' academic achievements, little is known about how peers influence each other's noncognitive outcomes. We attempt to fill this gap in the literature by utilizing unique data on individual students' mental health, social acclimation, and misbehavior in school. These outcomes are not only interesting for providing a complete picture of students development through schooling, but also for being good predictors of people's long-run wellbeing. Since Jencks et al. (1979), studies have documented the importance of noncognitive skills in explaining academic achievement, labor market success, and other significant life outcomes (Heckman and Rubinstein 2001; Heckman, Pinto and Savelyev 2013; Segal 2013; Bertrand and Pan 2013). We therefore attempt to push the boundary of student outcomes and explicitly consider students' noncognitive skills as a form of output of the educational production process.

A challenge in uncovering peer effects at school is the nonrandom grouping of students. Students with similar backgrounds or characteristics tend to associate with each other, and peer groups tend to be self-selected.[2] For our research question, if there are unobserved

[1] See Epple and Romano (2011) for an extensive review of the literature.

[2] Manski (1993) documents three types of effects that can generate similar peer outcomes: (1) correlated effects arise when individuals with similar backgrounds self-select into the same group; (2) exogenous effects

characteristics of schools and students that are correlated with both gender composition in the classroom and individual students' outcomes, the estimation of gender peer effects would be biased. To deal with the identification problem, researchers exploit cross-cohort variation (Hoxby 2000; Gould, Lavy and Daniele Paserman 2009; Carrell, Fullerton and West 2009; Lavy and Schlosser 2011; Black, Devereux and Salvanes 2013), or use random assignment (Sacerdote 2001; Zimmerman 2003; Carrell, Fullerton and West 2009; Kremer, Duflo and Dupas 2011; Shue 2013). We rely on unique information on classroom assignments, obtained from the survey questionnaire, and focus on middle schools in which students are randomly assigned to classrooms. Individuals in our refined sample cannot self-select into classrooms, and those assigned to the same classroom stay together for learning and extracurricular activities throughout the three years of middle school.

We use the China Education Panel Survey 2014 (CEPS 2014), a nationally representative survey of middle-school students and teachers, to estimate how the proportion of female classmates affects students' academic and noncognitive outcomes. We restrict our sample to schools that randomly place students in classrooms, and conduct balancing tests of students' predetermined characteristics to verify the identification assumption. In the regression analysis, we further control for school fixed effects to remove any school-level factors that influence students' achievements. The main outcome variables include students' test scores, which are obtained from school administrators, and noncognitive outcomes, which are obtained from the student survey. The noncognitive outcomes measure students' mental stress, social acclimation and general satisfaction at school, and disciplinary problems.

The results suggest that having higher proportion of female peers in the class positively affects students' academic and noncognitive outcomes. Specifically, a 10 percentage point increase in the proportion of female classmates raises student test scores by 4.77% of a standard deviation. The presence of more female peers also improves the average student's mental status, social acclimation and adaptation in school, and reduces the probability that he/she misbehaves. In addition, the results suggest that the positive effects on test scores are mostly driven by boys, while the effects on noncognitive outcomes are observed for boys and girls.

We further explore the potential mechanisms through which having more female peers in the class improves test scores and noncognitive outcomes. We find support for the channel through improved classroom environment, student-teacher interactions, and teachers' job satisfaction. When there are more female students in the class, students feel that the envi-

arise when individuals' predetermined characteristics affect their peers' outcomes; and (3) endogenous effects arise when individuals' outcomes directly affect their peers' outcomes. Since we are interested in the effects of gender, which is fixed, the endogenous effects are not applicable here. The focus of our identification strategy is to separate the exogenous effects from the correlated effects.

ronment is friendlier and more satisfying; they receive more feedbacks and more praises from teachers, and teachers report a higher level of job satisfaction. These changes associated with gender composition may be attributed to the observed benefits in students' learning and behavioral outcomes. However, the results do not support the hypothesis about ability-based spillover from female students. Instead, it is the subjects where girls have the lowest advantage that we identify the greatest positive gender peer effects.

This study complements and extends existing work on gender composition in schools. First, we add new evidence on the positive effects of the presence of having more female peers on students' academic achievements. Earlier studies on gender peer effects mainly compare outcomes from single-sex and co-educational schools, and the results are mixed: some studies suggest that single-sex schools might be beneficial, while others suggest there is no difference.[3] It is difficult to interpret these findings, since most of the studies cannot isolate the causal effect of peers on student performance.

More recent studies on gender peer effects use experimental or quasi-experimental strategies to deal with potential endogenous sorting and self-selection. Hoxby (2000) uses the idiosyncratic variation in adjacent cohorts in the same schools to examine gender and race peer effects in public elementary schools in Texas. Lavy and Schlosser (2011) exploit variations in gender composition at the school-grade level, and explore the gender peer effects and mechanisms in elementary, middle, and high schools in Israel. Black, Devereux and Salvanes (2013) exploit the idiosyncratic variation in cohort composition in middle schools in Norway to examine the effects of peer composition. Lu and Anderson (2015) experimentally vary the gender of peers sitting nearby in a secondary school in China. Most of these studies find positive effects of having more female peers on students' academic outcomes (Hoxby 2000; Lavy and Schlosser 2011; Lu and Anderson 2015).

We contribute to the literature by using a national sample of middle schools and focusing on random student assignment. Doing so allows us to gain a more representative and general perspective and at the same time address the endogenous grouping problem. Previous studies that have used national surveys typically lack the information on how students are assigned to classes (Ehrenberg, Goldhaber and Brewer 1995; Dee 2005; Jepsen 2005), and studies that use experimental designs usually can only focus on one school or district, and thus have limited generalizability.

Second, this study complements the existing work by documenting the gender peer effects on noncognitive outcomes. Individual students' noncognitive outcomes could explain their academic achievements and long-run labor market success and well-beings. Lavy and

[3]See Morse (1998) and Harwarth, DeBra and Maline (1997) for reviews of earlier studies on gender peer effects.

Schlosser (2011), the closest to this study, also examine noncognitive outcomes, such as violence and disruptive behavior, but their focus is on students' perceptions and general feelings about the classroom environment. Our focus here is on on individual-level mental health, social acclimation, and misbehavior.

Our findings also build on the broad literature on the importance of the school and residence environments for students' noncognitive and behavioral outcomes. Previous studies have documented the effects of school choice (Cullen, Jacob and Levitt 2006; Angrist, Bettinger and Kremer 2006; Figlio and Ludwig 2000; Lavy 2010), school social networks (Lavy and Sand 2015), and neighborhood composition (Gibbons, Silva and Weinhardt 2013) on students' outcomes, such as social acclimation, anti-social behaviors, overall happiness, and behavioral problems and discruptions in school. We provide evidence of yet another school factor that may influence students' noncognitive outcomes; therefore, policy makers should account for gender composition more carefully in designing the optimal learning environment.

Finally, our results contribute to the understanding of peer effects in the educational production function. In this literature, the definition of peers varies by context, including peer cohorts within the same school (Angrist and Lang 2004; Arcidiacono and Nicholson 2005; Ammermueller and Pischke 2009; Gould, Lavy and Daniele Paserman 2009), roommates in college dorms (Sacerdote 2001; Zimmerman 2003) and peer groups in military academies (Carrell, Fullerton and West 2009; Lyle 2007). Peer effects have been found in several dimensions. Along the ability dimension, previous studies have found that peers' abilities have positive effects on student achievement (see, e.g., Sacerdote 2001; Zimmerman 2003; Ammermueller and Pischke 2009). Along the race and social background dimension, Angrist and Lang (2004) evaluate effect of the Metco Program, which sent minority students to schools in affluent suburbs of Boston; and they find modest and short-lived peer effects. Gould, Lavy and Daniele Paserman (2009) show that the overall presence of immigrants in a grade adversely affect students' academic achievement. This study focuses on peers in the classroom and our findings have implications for the benefits of co-educational schooling and the optimal grouping of students.

2 Data and Variables

The empirical analysis draws on data from the 2014 China Education Panel Survey (CEPS), a nationally representative survey conducted by the China National Survey and Research Center. The CEPS is the first and largest nationally representative educational survey in China, including 19,487 students in 438 classes in 112 middle schools. The survey applies a stratified sampling design such that four middle schools and four classrooms within each school are chosen to represent a given county or city district. The final merged data set includes questionnaire information from students, their parents, teachers and schools principals.

The measure of peer gender is the proportion of female peers in the same class. Typically, in middle schools in China, students are assigned to classes at the beginning of the 7th grade and take the same courses throughout their three years in middle school. During a regular school day, students remain in the same classroom all day and teachers come to deliver lectures in each subject. In addition of taking courses together, students in the same class also participate in a variety of exercises and activities together, such as self-study sessions, sports events, and field trips. As a result, students in the same class have intensive interactions with academic and non-academic purposes.

Student outcome variables include academic performance and noncognitive outcomes. Students' academic performance is measured by their exam scores in the core courses, which are provided by the school administration offices. We focus on the core courses because these subjects—Chinese, mathematics and English—are the main components in the standards tests for admission into senior high schools, and are mandatory for all middle schools. Within a school, teachers teaching the same course use an identical syllabus and give the same exams during a common testing period. Exams are graded in a rigorous and consistent manner. During the grading process, each student's name, class, and ID are hidden from the graders. Within a grade in the same school, teachers divide the grading work so that the same question is typically graded by the same teacher with a consistent standard. Test scores in the core courses are therefore a consistent measure of academic achievement across students in the same grade in the same school. We supplement the test scores with students' self-assessment academic performance scores. In the survey, students are asked to report to what extent they have difficulties in learning each subject, with a scale of 1 (have lots of difficulties in learning) to 4 (have no difficulties in learning). To facilitate interpretation, we normalize the test scores and self-assessment scores within the subject, grade, and school to have a mean of zero and standard deviation of one.[4]

Measures of noncognitive outcomes are obtained from students' responses to 10 survey items. Four questions are related to students' mental stress, and in particular, ask whether during the previous seven days the students felt (1) depressed, (2) blue, (3) unhappy, and (4) that life is meaningless. Four questions ask about students' social acclimation and general satisfaction at school, in particular, (5) whether they feel that school life is boring, (6) whether they feel confident about their future, (7) how often they go to museums, zoos or

[4]To facilitate interpretation, we reverse the scale of the self-assessment score in questionnaire, so that 1 is the lowest score and 4 is the highest.

science parks with classmates, and (8) how often they go to movies, plays or sporting events with classmates. Finally, two questions ask about students' misbehavior, in particular, (9) how often they are late for class, and (10) how often they skip class.[5] To facilitate interpretation, we normalize each outcome variable so that it has a mean of zero and standard deviation of one.

Table 1 presents the summary statistics of our main variables—students' academic and noncognitive outcomes, own and peers' gender, and students' basic demographics. Students' raw test scores (un-normalized) have a mean of 81.209, and their self-assessment scores (reported on a scale of 1 to 4) have a mean of 2.472. Both objective and subjective scores of academic performance have large standard deviations, suggesting wide dispersion among students. Students' own and peers' gender show that about 48% of the students in the sample are female and, not surprisingly, this is the average proportion of female peers a given student has.

[Insert Table 1 here]

We also present students' predetermined characteristics for the purpose of a balance test and further control in the main specification. These variables are *Age*, *Minority* (1 if a student belongs to a minority ethnicity group and 0 otherwise), *Local Residence* (1 if a student is from the local area and 0 otherwise), *Only Child* (1 if a student is the only child in the family and 0 otherwise), *Attend Kindergarten* (1 if a student attended kindergarten and 0 otherwise), *Repeat in Primary School* (1 if a student repeated a grade in the primary school), *Mother Education* (mother's years of schooling), and *Father Education* (father's years of schooling).

3 Estimation Strategy

To investigate how peers' gender composition affects students' cognitive and noncognitive development, we implement the linear-in-means model, which has been widely adopted in the literature (e.g., Sacerdote 2001; Guryan, Kroft and Notowidigdo 2009; Lu and Anderson 2015). We use the following regression model:

$$Y_{ics} = \alpha + \beta_1 Peerfem_{-ics} + \beta_2 Female_{ics} + \phi X'_{ics} + \tau_g + \lambda_s + \varepsilon_{ics}, \tag{1}$$

[5]For items (1) to (4), respondents are asked to report the frequency of incidents on a scale from 1 (never) to 5 (always). For items (5), (6), (10), and (11), they were asked to rate how much they agree with the statements on a scale from 1 (strongly disagree) to 4 (strongly agree). For items (7) and (8), they were asked to rate how much they agree with the statement on scale from 1 (never) to 6 (always).

where Y_{ics} are the outcome measures for student i in class c in school s; $Peerfem_{-ics}$ captures the proportion of female students in student i's class excluding student i; $Female_{ics}$ is an indicator for whether the student i is female; and ε_{ics} is the error term. We cluster the standard errors at the class level, accounting for correlation in outcomes for students in the same class.

The coefficient of interest is β_1, which captures the gender peer effects on academic outcomes and noncognitive skills. The unbiased estimation requires that conditional on all the controls in the equation, our regressor of interest $Peerfem_{-ics}$ is uncorrelated with the error term ε_{ics}. A possible threat to the identifying assumption is that students may select into classes by unobservables; hence, β_1 may reflect the sorting of students with certain characteristics rather than the gender peer effect.

To address this identification concern, we focus on schools in the data set that randomly assign students to classes, similar to the strategies used in Sacerdote (2001), Zimmerman (2003), Carrell, Fullerton and West (2009) and Shue (2013). We describe the institutional background of class assignment and present validity checks on random assignment in the following two subsections. With students randomly assigned to classes, β_1 is an unbiased estimator of the gender peer effects. Furthermore, as random class assignment is conducted within schools and school choices may not be random, we include school fixed effects λ_s in the regressions, controlling for all school-level factors that may influence the students' school choice decisions. We also include grade fixed effects τ_g to account for differences between different grades. In addition to the baseline estimates, we further present results that control for students' predetermined characteristics X_{ics} (i.e., values determined before the class assignment), which can improve the balance between classes and our estimation efficiency.

3.1 Class Assignment and Estimation Sample

Our key research question concerns the gender peer effects on student outcomes. It is therefore critical to understand how students are assigned to classrooms and how classmates interact. In China, middle school students are assigned to classes at the beginning of 7th grade. Once assigned, students in the same class typically remain together throughout the three years of study. They take the same subjects and are lectured by the same teachers. In addition to studying together, students attend and compete in extracurricular events—sport, arts and social activities—representing their class. Therefore, classmates tend to interact closely and intensively, and influence each other in the learning and social dimensions.

The way students are matched to classrooms and therefore varying peer composition is critical to our estimation of gender peer effects. Middle schools in China adopt different strategies to assign students to classes. In some schools, prior to beginning their first academic year, students take placement exams and their scores are used to assign them to classes. Other schools sort students by residential status, and students are placed in classes for local residents or migrants. These types of assignments are nonrandom and therefore may bias our estimation of gender peer effects.

More recently, a large and increasing number of schools have employed random assignment rules to place students in classes. This approach is heavily promoted by the Ministry of Education to ensure equal and fair opportunity for all students during compulsory education (up to the 9th grade). Typically, randomization of students into classes is conducted by a computer program, which can also account for balance between classrooms in terms of class size, the proportion of migrants and other important features. Some schools also try to balance the composition of student ability, and therefore use students' scores and rankings in a placement exam as the balancing factor when running the randomization program. Another commonly observed approach, especially when enrollment is small, is to invite parents to draw lots to determine which classes their children will join. Once student assignments are completed, teachers also draw lots to determine which classes they will teach and manage.

We use principal and teacher responses in the survey to identify schools that randomly assign their students. Such identification was challenging for previous studies, because assignment information is usually missing in other national-wide surveys (e.g., Ehrenberg, Goldhaber and Brewer 1995; Dee 2005; Jepsen 2005). We consider that a school conducts random assignment if its procedure meets all the following conditions: (1) the school principal reports that students are randomly assigned to classes; (2) after students are assigned to classes at the beginning of their 7th grade, the school does not rearrange their classes for grades 8 and 9; and (3) all the head teachers in the same grade report that students in their grade are not assigned by test scores.[6] In the 2014 CEPS data, 59.8% of the schools use random assignment by our criteria, resulting in an estimation sample of 8,988 students in 208 classes in 67 schools. Each student in our sample is therefore randomly placed before enrollment and stays in the same classroom throughout his or her three years of study. It ensures that there is no self-selection of students to classes or certain peers, and particularly, that students assignment is random with respect to gender composition in the classroom.

[6]The criteria are based on reports in the principal and teacher questionnaire. For the first condition, all school principals are asked to report, in a multiple choice question, which assignment rule they adopted to place students: (a) based on pre-enrollment test scores, (b) based on students' residential status, (c) random assignment, or (d) based on other factors. We restrict our sample to schools that use assignment method (c). Second, the same principals are asked whether their schools rearrange classes for grades 8 and 9, and we rule out those that report "yes". Last, each head teacher was asked whether students in his or her grade are assigned by test scores, and we drop the whole grade if any head teacher in the grade reported "yes".

3.2 Verifying Random Class Assignment

To verify that students are randomly assigned to classes in our regression sample, we follow the standard methods and conduct balancing tests of students' characteristics in the baseline, i.e., variables determined before the class assignment. We use the baseline specification (1) to check whether students' predetermined characteristics—age, ethnicity, residential status, whether they have any siblings, kindergarten and primary school background, and parents' education—are balanced across groups with different percentages of female peers. If class assignment is indeed random, students having different proportions of female peers in their classes should be similar in these observed characteristics. As shown in Table 2, all the estimated coefficients of $Peerfem_{-ics}$ are statistically insignificant, except for *Mother Education*. The magnitudes of all the coefficients are small relative to the sample mean. For example, the coefficient for mother's education is 1.290, suggesting that an one-standard-deviation increase in $Peerfem_{-ics}$ (0.085) is associated with an increase of $0.085 \times 1.290 = 0.110$ years of schooling for a student's mother, which is only about 1 percent of the sample mean. These results suggest that students with different percentages of female peers are overall balanced in their baseline characteristics.

[Insert Table 2 here]

In checking the balance of students' own gender, we follow the approach used by Guryan, Kroft and Notowidigdo (2009) to account for potential bias caused by sampling peers without replacement. Specifically, as a student cannot be assigned to herself, sampling of peers is done without replacement. In our setting, one direct consequence is that the peers for a female students are chosen from a group with fewer female than the peers for a male student in the same class. Guryan, Kroft and Notowidigdo (2009) suggest a solution by additionally controlling for the mean of the sampling pool, that is, the proportion of all female peers in the same grade in the same school excluding student i. The estimation equation is

$$Female_{ics} = \alpha + \pi_1 Peerfem_{-ics} + \pi_2 Peerfem_{-igs} + \gamma_{gs} + \varepsilon_{ics}, \tag{2}$$

where $Peerfem_{-igs}$ measures the proportion of female peers in the same grade in the same school excluding student i; and γ_{gs} is the school-grade fixed effects. The estimation results are reported in the last row of Table 2. We find that the coefficient of $Peerfem_{ics}$ is small in magnitude and statistically insignificant. This result lends further support to our identification assumption that students were randomly assigned to classes in a school.

4 Main Results

4.1 Gender Peer Effects on Academic Performance

Table 3 reports the estimated gender peer effects on students' academic outcomes, measured by test scores and self-assessment performance. We pooled three subjects (Chinese, math and English) together and include subject fixed effects, school fixed effects, and grade fixed effects. To facilitate interpretation, all the variables are normalized by school, grade and subject to have a mean of zero and a standard deviation of one. We report the results without controlling for student predetermined variables in the odd columns, and those with these controls in the even columns.

[Insert Table 3 here]

Columns (1)-(2) and columns (3)-(4) report the effects of the proportion of female peers on students' test scores and self-assessment scores, respectively. The results show that when students have a higher proportion of female peers in the class, they achieve higher test scores. Interestingly, the gender peer effects on students' self-assessment performance are small and statistically insignificant. These findings imply that having more female classmates improves academic performance but does not significantly increase the students' confidence, as measured by self-assessment.

To appreciate the economic significance of the effect on test scores, we use the more conservative estimate in column (2), which controls for student characteristics. The coefficient suggests that a 10 percentage point increase in the proportion of female classmates raises a student's test score by 4.77% of a standard deviation. The magnitude of the effect is comparable to the estimates in previous studies. For instance, Lavy and Schlosser (2011) show that a 20 percentage point increase in the share of female students raises 8th graders' by 6 to 8 percent of a standard deviation.

4.2 Gender Peer Effects on Noncognitive Outcomes

To examine the gender peer effects on students' noncognitive outcomes, we focus on 10 items in the student questionnaire that relate to their mental stress, social acclimation and general satisfaction at school, and disciplinary problems. All outcome variables are normalized to have a mean of zero and a standard deviation of one. By definition, lower scores in mental status and disciplinary problems and higher scores in social acclimation and general satisfaction point to improved noncognitive outcomes. For each group, we first report the effects on the outcomes themselves (for example, whether the student felt despressed in the previous seven days). Then, following Kling, Liebman and Katz (2007), we present the average effect sizes (AES) to assess the overall effects on each category (for example, the average effect across various measures of mental stress).[7]

Table 4, panel A, presents the estimated effects on students' mental health. The estimated AES in column (1) is negative and significant, suggesting that, overall, having a higher proportion of female peers in the class reduces a student's mental stress. Columns (2)-(5) report the effects on four aspects of mental conditions separately. All the coefficients are negative, consistent with the overall pattern, despite that the effects on the likelihood of feeling depressed or blue are imprecisely estimated. Controlling for student-level characteristics in columns (7)-(10) yields results of similar pattern.

[Insert Table 4 here]

Table 4, panel B, reports the outcomes on social acclimation and general satisfaction at school. Overall, we find a positive effect from having more female classmates on students' outcomes along this dimension. The effects on the four individual measures of social acclimation and general satisfaction at school are all positive and statistically significant, except for the likelihood to engage in public enrichment with their classmates.

Last, Table 4, panel C, reports the estimated effects on students' disciplinary problems. The estimated AES is negative and statistically significant, indicating that having more female classmates helps to reduce disciplinary problems. Columns (2) and (3) further show that having more female classmates significantly reduces the chances that students are late for school or skip classes. All the effects remain negative and significant after controlling for student-level characteristics.

Overall, the abovementioned results provide consistent evidence that having a higher proportion of female peers in the class significantly improves students' mental status and social acclimation, and reduces the probability that they misbehave.

4.3 Robustness Checks

We conduct two sets of robustness checks in this section: more flexible specifications that allow for nonlinear effects of peer gender, and the test of sample attrition.

Nonlinear effects. Our baseline estimation specification applies a linear-in-means model, capturing the linear effects of the proportion of female peers. Here, we consider the possibility

[7] Specifically, we define the average effect size (AES) of the proportion of female peers on category c as $AES = \frac{1}{n_c}\sum_{n=1}^{n_c} e_{kc}/\sigma_{kc}$, where n_c is the number of outcomes in category c, e_{kc} is the estimated effect for outcome k of category c, and σ_{kc} is the standard deviation of the outcome variables.

that the effects of female peers could be nonlinear as in the works by Hoxby (2000), Lavy and Schlosser (2011) and Black, Devereux and Salvanes (2013). Specifically, we modify specification (1) by including dummy variables for four categories of female ratio: (1) 20% to 40%, (2) 40% to 60%, (3) 60% to 80%, and (4) 80% to 100%. The omitted category is a dummy variable indicating the proportion of female peer below 20%.

[Insert Table A1 here]

The results, as shown in Appendix Table A1, show some extent of nonlinearity and there appears to be a general upward trend in most of the coefficients of the proportion of females. In particular, several estimates (such as self-assessed score, sociability measurements, and late for school) suggest that the gain from having more female peers is mostly driven by the presence or absence of a very high proportion of female peer. This finding is in line with Hoxby (2000) in which the impacts are highest when female students are a majority of the class, and Lavy and Schlosser (2011) in which the impacts are highest when the proportion of females exceeds 58%.

Sample attrition. One concern on sample attrition is that, our estimates of gender peer effects may be biased if peer's gender affects the likelihood of missing values in students' outcome variables. For example, having more female peers may make some students less willing to report disciplinary problems. To address this concern, we generate a attrition dummy indicating missing values for each outcome of interest, then regress the attrition dummy on peer gender, student gender, and school and grade fixed effects. As shown in Appendix Table A2, the coefficients on peer gender (proportion of female peer) are all closed to zero and not statistically significant, implying that our results are unlikely to be driven by sample attrition.

[Insert Table A2 here]

5 Mechanisms

The aforementioned estimates show positive effects of female peers in the classroom on students' academic performance and noncognitive outcomes. In this section, we explore the potential channels through which having more female peers benefits a student. In particular, we focus on four potential channels: positive spillover from female peers, improved classroom environment, better teacher-student interactions, and improved teachers' job satisfaction.

5.1 Female Students Spillovers

It has been well recognized and documented that girls outperform boys in middle schools, especially in subjects such as reading and language, and in behavior (Dee 2007; Bertrand and Pan 2013). Hence, when more female students—who are likely to be high-achieving in academic tests—are present in class, other students could benefit from a positive spillover. In other words, female students on average exhibit higher ability, and the effects of having more female students in the classroom stem from the spillovers from those high-ability students.

We argue that if the gender peer effects are primarily driven by spillover from high-achieving female students, then the effect is likely to concentrate on subjects and areas where girls have greater advantage over boys. We examine the three core subjects separately to exploit the variation in female advatange. As shown in Table 5, the coefficients of *Female Student* suggest that girls' advantage is larger in Chinese and English and much smaller in math. This is consistent with the evidence in the literature in that compare with math and science, girls have comparative advantage in reading and language (Dee 2007; Fryer and Levitt 2010). However, the coefficients of *Proportion female peer*—our estimates of the gender peer effects—reveal that having more female peers yields greatest benefits on a student's math test scores, and small and statistically insignificant effects on Chinese test scores.

[Insert Table 5 here]

The contrast direction of the gender gap and gender peer effects does not support the spillover hypothesis. Indeed, similar contrast has been documented in exisiting studies. For instance, Hoxby (2000) show that although girls and boys perform equally well in math, the presence of more female peers increases the math scores of both female and male students. Lavy and Schlosser (2011) find positive effects from having more female peers in subjects that girls achieve lower scores than boys. So we claim that the spillover channel is unlikely to explain the positive effects of having more female students in the class.

5.2 Classroom Environment

A second possible mechanism is that female students make the classroom environment friendlier and more satisfying, leading to better academic and noncognitive behavioral outcomes for all students. To investigate this possible channel, we use the following two items: (1) "I feel that my classmates are friendly to me." and (2) "I feel that our classroom has a satisfying atmosphere." The students are asked to rate to what extent they agree with the statements on a scale from 1 (strongly disagree) to 4 (strongly agree). We normalize the responses with a mean of zero and a standard deviation of one to fit regression equation (1).

The results, as shown in Table 6, support the hypothesis. We observe that when more female students are present in the class, on average, students feel learning environment friendlier and the atmosphere more satisfying. The effects are statistically significant. It is possible that the improved classroom environment may make learning more enjoyable and efficient, thus benefit students' achievement on tests. Friendlier environment can also facilitate students social interaction so that they are well-adjusted among school peers. Overall, we find evidence supporting the channel that more female students in the class is associated with a better learning environment for the students.

[Insert Table 6 here]

5.3 Teacher-Student Interaction

The third mechanism we examine is the interaction between students and their teachers. The gender composition in the classroom may influence teachers' behavior toward students and therefore students' learning outcomes. To assess the relevance of this channel, we use four items in the student questionnaire that describe their interaction with teachers: (1) "the head teacher always praises me;"(2) "the head teacher always criticizes me;"(3) "the subject teachers always ask me to answer questions in class;"and (4) "The subject teachers always praise me in class."The students are asked to rate to what extent they agree with the statements on a scale from 1 (strongly disagree) to 4 (strongly agree), and we normalize the responses with a mean of zero and a standard deviation of one.

Table 7 reports the estimation results, which support improved teacher-student interactions when more female students are present in class. Results show that when there is a higher proportion of female peers in the class, a student is likely to receive more praise and less criticism from the head teacher; he or she is also more likely to be questioned in class and receive praise from subject teachers. Overall, these results show that there is more proactive and positive interaction between the teacher and students when there are more female students in the classroom, lending support to the machanism through better teacher-student interaction.

[Insert Table 7 here]

5.4 Teacher's Job Satisfaction

Last, we analyze whether more female students in the class affects teachers' work satisfaction and evaluation of the classroom environment. Teacher-level factors may influence their motivation and the way they teach and manage the class. To this end, we focus on three items from the teachers' questionnaire: (1) "I'm satisfied with being a teacher;"(2) "I feel that the school atmosphere is good;"and (3) "The students in class respect me."The teachers are asked to rate to what extent they agree with the statements on a scale from 1 (strongly disagree) to 5 (strongly agree), and we normalize the responses with a mean of zero and a standard deviation of one.

Table 8 report the estimation results, revealing that teachers feel more satisfied with the job and consider the classroom environment better when more female students are present in the class. Although the effects are not statistically significant, the positive association is consistent with the hypothesis that gender composition in the class influence teachers' assessment and possibly their teaching behavior.

[Insert Table 8 here]

Overall, the findings around possible mechanisms suggest that the positive effects of having more female peers in the class are not entirely generated by the spillover from girls' better performance. Instead, we find evidence that female students improves the classroom environment, inter-students and teacher-student relationships, and teachers' job satisfaction. The evidence is in line with Lavy and Schlosser (2011), who focus on the Israel national education data and find that an increased proportion of female peers reduces the level of disciplinary problems, improves inter-students and teacher-student relationships, and reduces teachers' fatigue.

Unfolding the mechanisms is important for policy design. To the extent that the number of female students in a school is fixed, the benefit from having more female peers in a class could be offset by the cost of having fewer female peers in another class. The source of the peer effects provides more practical and low-cost opportunities to improve student outcomes. For example, teachers can exert more effort in facilitating student interactions and providing more feedback if there are fewer female students than desired. The assignment of teachers to classes can also take into account the gender composition: teachers who interact more with students might compensate for the lower proportion of female students in the classroom.

6 Gender Differences in Gender Peer Effects

In this section, we explore how female and male students are affected by peers' gender in different ways. Table 9 reports the gender peer effects on test scores and self-assessment scores for boys and girls, respectively. As for test scores, we find small and insignificant impact of peers' gender on girls but large and significant effects on boys. On average, a 10 percentage point increase in the proportion of female classmates is associated with a 0.7% standard deviation increase in girls' test scores but a 9.6% increase in boys' test scores. The difference between the coefficients is statistically significant. Neither boys' or girls' self-assessment scores appear to be responsive to the share of female peers. The coefficients for each gender group are small and statistically insignificant, and the difference between the effects on girls versus boys is also insignificant.

[Insert Table 9 here]

Table 10 reports the heterogeneous effects on noncognitive skills by student gender. The benefit of more female peers on reducing mental stress is more pronounced among girls than boys, while the benefit of reducing disciplinary problems is stronger for boys. It is possible that girls and boys differ in the sensitivity of their emotion and behavior to environments. Girls' mental status might be more sensitive to their surroundings and peers than boys, while boys' behavior are more likely to be corrected by environmental influences. The effects on social acclimation and satisfaction are similar, with slightly larger magnitudes for boys.

[Insert Table 10 here]

7 Conclusion

This paper uses a nationally representative survey of middle school students to investigate gender peer effects on students' academic performance and various noncognitive outcomes. Utilizing information about classroom assignment within schools, we are able to restrict the sample to schools that randomly assign students in classrooms and therefore estimate the causal relationship between peers' gender and student outcomes.

The results show that having a higher proportion of female peers in the class significantly raises students' test scores, reduces their mental stress, improves social acclimation and general satisfaction at school, and mitigates disciplinary problems. Exploring the potential mechanisms through which peers' gender plays a role, we find evidence of improved classroom environment, inter-students and teacher-student interactions, and teachers' level of job satisfaction.

The findings have several contributions and policy implications. First, while most of the literature focuses on the effects of the school environment on students' academic outcomes, our study provides more evidence on the impact on students' noncognitive and behavioral outcomes, which have been documented as important factors in explaining academic achievement, labor market success, and other significant life outcomes (Heckman and Rubinstein 2001; Heckman, Pinto and Savelyev 2013; Segal 2013; Bertrand and Pan 2013). Second, understanding the mechanisms behind gender peer effects provides relatively low-cost methods for educational policy design to improve students' test scores and noncognitive outcomes. For example, to compensate for the small share of female students in certain classes, schools could allocate more experienced and patient teachers to classes, or provide orientations and class activities to facilitate students' cooperation and interactions.

References

Ammermueller, Andreas and Jörn-Steffen Pischke, "Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study," *Journal of Labor Economics*, 2009, *27* (3), 315–348.

Angrist, Joshua D and Kevin Lang, "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," *The American Economic Review*, 2004, *94* (5), 1613–1634.

Angrist, Joshua, Eric Bettinger, and Michael Kremer, "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia," *The American Economic Review*, 2006, *96* (3), 847–862.

Arcidiacono, Peter and Sean Nicholson, "Peer Effects in Medical School," *Journal of Public Economics*, 2005, *89* (2), 327–350.

Bertrand, Marianne and Jessica Pan, "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior," *American Economic Journal: Applied Economics*, 2013, *5* (1), 32–64.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "Under Pressure? The Effect of Peers on Outcomes of Young Adults," *Journal of Labor Economics*, 2013, *31* (1), 119–153.

Carrell, Scott, Richard L Fullerton, and James West, "Does Your Cohort Matter? Measuring Peer Effects in College Achievement," *Journal of Labor Economics*, 2009, *27* (3), 439–464.

Cullen, Julie Berry, Brian A Jacob, and Steven Levitt, "The Effect of School Choice on Participants: Evidence from Randomized Lotteries," *Econometrica*, 2006, *74* (5), 1191–1230.

Dee, Thomas S, "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?," *The American Economic Review*, 2005, *95* (2), 158–165.

_ , "Teachers and the Gender Gaps in Student Achievement," *Journal of Human Resources*, 2007, *42* (3), 528–554.

Ehrenberg, Ronald G, Daniel D Goldhaber, and Dominic J Brewer, "Do Teachers' Race, Gender, and Ethnicity Matter? Evidence from the National Educational Longitudinal Study of 1988," *Industrial & Labor Relations Review*, 1995, *48* (3), 547–561.

Epple, Dennis and Richard Romano, "Peer Effects in Education: A Survey of the Theory and Evidence," *Handbook of Social Economics*, 2011, *1* (11), 1053–1163.

Figlio, David and Jens Ludwig, "Sex, Drugs, and Catholic Schools: Private Schooling and Non-Market Adolescent Behaviors," Technical Report, National Bureau of Economic Research 2000.

Fryer, Roland G and Steven D Levitt, "An Empirical Analysis of the Gender Gap in Mathematics," *American Economic Journal: Applied Economics*, 2010, *2* (2), 210–240.

Gibbons, Stephen, Olmo Silva, and Felix Weinhardt, "Everybody Needs Good Neighbours? Evidence from Students Outcomes in England," *The Economic Journal*, 2013, *123* (571), 831–874.

Gould, Eric D, Victor Lavy, and M Daniele Paserman, "Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence," *The Economic Journal*, 2009, *119* (540), 1243–1269.

Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo, "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," *American Economic Journal: Applied Economics*, 2009, *1* (4), 34–68.

Harwarth, Irene, Elizabeth DeBra, and Mindi Maline, *Women's Colleges in the United States: History, Issues and Challenges*, DIANE Publishing, 1997.

Heckman, James J and Yona Rubinstein, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *The American Economic Review*, 2001, *91* (2), 145–149.

Heckman, James, Rodrigo Pinto, and Peter Savelyev, "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes," *The American Economic Review*, 2013, *103* (6), 1–35.

Hoxby, Caroline, "Peer Effects in the Classroom: Learning from Gender and Race Variation," Technical Report, National Bureau of Economic Research 2000.

Jencks, Christopher et al., "Who Gets Ahead? The Determinants of Economic Success in America," 1979.

Jepsen, Christopher, "Teacher Characteristics and Student Achievement: Evidence from Teacher Surveys," *Journal of Urban Economics*, 2005, *57* (2), 302–319.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Kremer, Michael, Esther Duflo, and Pascaline Dupas, "Peer Effects, Teacher Incentives, and the Impact of Tracking," *American Economic Review*, 2011, *101* (5).

Lavy, Victor, "Effects of Free Choice among Public Schools," *The Review of Economic Studies*, 2010, *77* (3), 1164–1191.

_ **and Analia Schlosser**, "Mechanisms and Impacts of Gender Peer Effects at School," *American Economic Journal: Applied Economics*, 2011, *3* (2), 1–33.

_ **and Edith Sand**, "The Effect of Social Networks on Students Academic and Non-Cognitive Behavioral Outcomes: Evidence from Conditional Random Assignment of Friends in School," *University of Warwick and Hebrew University Working Paper*, 2015.

Lu, Fangwen and Michael L Anderson, "Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups," *Journal of Labor Economics*, 2015, *33* (1), 91–122.

Lyle, David S, "Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point," *The Review of Economics and Statistics*, 2007, *89* (2), 289–299.

Manski, Charles F, "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

Morse, Susan, *Separated by Sex: A Critical Look at Single-Sex Education for Girls*, Amer Assn of Univ Women, 1998.

Sacerdote, Bruce, "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *The Quarterly Journal of Economics*, 2001, *116* (2), 681–704.

Segal, Carmit, "Misbehavior, Education, and Labor Market Outcomes," *Journal of the European Economic Association*, 2013, *11* (4), 743–779.

Shue, Kelly, "Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers," *Review of Financial Studies*, 2013, *26* (6), 1401–1442.

Zimmerman, David J, "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment," *Review of Economics and Statistics*, 2003, *85* (1), 9–23.

Table 1. Summary Statistics

	(1)	(2)	(3)
	Mean	Standard deviation	Observation
Outcome variable: Academic outcomes			
Test score	81.209	28.400	26209
Self-assessment score	2.472	0.916	26746
Outcome variable: Noncognitive outcomes			
Depressed	2.238	1.000	8772
Blue	1.980	1.060	8743
Unhappy	2.277	1.047	8762
Pessimistic	1.747	1.072	8734
School life is fulfilling	3.375	0.861	8852
Confident about future	3.256	0.717	8924
Social activity: Public enrichment	2.025	1.037	8686
Social activity: Private recreation	2.437	1.283	8653
Being late for school	1.246	0.618	8931
Skip class	1.091	0.427	8924
Regressors of interests			
Proportion of female peer	0.487	0.085	8910
Female student	0.487	0.500	8910
Predetermined variables			
Age	13.941	1.348	8815
Minority	0.110	0.313	8968
Local residence	0.804	0.397	8811
Only child	0.509	0.500	8986
Attend kindergarten	0.817	0.387	8912
Repeat in primary school	0.115	0.319	8988
Mother’s education	10.075	3.634	8966
Father’s education	10.727	3.284	8966

Table 2. Balancing Test

		(1)	(2)
	Sample mean	Simple OLS	With school fixed effects
Age	13.941	-0.174	-0.092
		(1.008)	(0.092)
Minority	0.110	-0.184	-0.052
		(0.305)	(0.035)
Local residence	0.804	0.298	-0.005
		(0.203)	(0.094)
Only child	0.509	0.592**	0.117
		(0.236)	(0.085)
Attend kindergarten before primary school	0.817	0.050	0.094
		(0.144)	(0.078)
Repeat in primary school	0.115	-0.340***	-0.015
		(0.130)	(0.051)
Mother's education	10.075	4.953**	1.290**
		(2.252)	(0.622)
Father's education	10.727	4.197**	*0.671*
		(1.676)	(0.613)
Female	0.487	0.000*	0.010
		(0.000)	(0.031)

Notes: In column (1) and column (2), each cell represents a separate regression, in which the dependent variable is the student's predetermined variable as listed above. See section 3.2 for detailed descriptions. Specifications in column (2) include school fixed effects and grade fixed effects. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Table 3. Gender Peer effect on Academic Outcomes

	Test score		Self-assessment score	
	(1)	(2)	(3)	(4)
Proportion female peer	0.493**	0.477**	0.103	0.046
	(0.247)	(0.226)	(0.161)	(0.154)
Female student	0.439***	0.414***	0.189***	0.174***
	(0.021)	(0.022)	(0.017)	(0.017)
Subject fixed effects	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes
Individual controls	No	Yes	No	Yes
Observations	25,992	24,735	26,520	25,277
R-squared	0.048	0.080	0.009	0.024

Notes: Test score and self-assessment score are normalized by subject, grade and school, to obtain a mean of zero and standard deviation of one. Individual control includes: student's age, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and repeated a grade in primary school. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Table 4. Gender Peer Effect on Noncognitive Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Mental stress										
	AES	Depressed	Blue	Unhappy	Pessimistic	AES	Depressed	Blue	Unhappy	Pessimistic
Proportion female peer	-0.336*	-0.109	-0.214	-0.535**	-0.526**	-0.276	-0.071	-0.113	-0.484**	-0.487**
	(0.194)	(0.231)	(0.209)	(0.214)	(0.219)	(0.185)	(0.221)	(0.202)	(0.207)	(0.215)
Female student		0.147***	-0.043*	0.034	-0.078***		0.155***	-0.038	0.044*	-0.071***
		(0.024)	(0.024)	(0.023)	(0.021)		(0.025)	(0.025)	(0.023)	(0.022)
Observations	8,607	8,696	8,667	8,687	8,658	8,213	8,294	8,268	8,285	8,260
R-squared		0.046	0.049	0.046	0.036		0.053	0.053	0.052	0.042
Panel B: Social acclimation and satisfaction										
				Social activity with classmates					Social activity with classmates	
	AES	Fulfilling of life	Confident abt future	Public enrichment	Private recreation	AES	Fulfilling of life	Confident abt future	Public enrichment	Private recreation
Proportion female peer	0.552***	0.532**	0.604***	0.338	0.723***	0.495***	0.514**	0.525***	0.258	0.654***
	(0.112)	(0.217)	(0.164)	(0.245)	(0.215)	(0.110)	(0.214)	(0.153)	(0.257)	(0.230)
Female student		0.094***	-0.013	0.050**	0.021		0.086***	-0.023	0.046**	0.021
		(0.021)	(0.022)	(0.021)	(0.021)		(0.022)	(0.022)	(0.021)	(0.021)
Observations	8,412	8,777	8,847	8,614	8,583	8,038	8,365	8,427	8,219	8,191
R-squared		0.076	0.074	0.205	0.209		0.083	0.084	0.219	0.224
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

Table 4. Gender Peer Effect on Noncognitive Outcomes (continued)

	Panel C: Disciplinary problems					
	(1)	(2)	(3)	(4)	(5)	(6)
	AES	Late for school	Skip class	AES	Late for school	Skip class
Proportion female peer	-0.327**	-0.252	-0.387**	-0.339**	-0.288*	-0.379**
	(0.143)	(0.168)	(0.151)	(0.149)	(0.174)	(0.160)
Female student		-0.090***	-0.100***		-0.089***	-0.097***
		(0.026)	(0.022)		(0.026)	(0.023)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	No	No	Yes	Yes	Yes
Observation	8,846	8,856	8,849	8,431	8,440	8,434
R-squared		0.073	0.073		0.075	0.076

Notes: Individual control includes: student's age, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and repeated a grade in primary school. Standard errors are clustered at class level and reported in parentheses.***significant at the 1% level, **5% level, *10% level.

Table 5. Mechanism: Spillover Effect

	(1)	(2)	(3)	(4)	(5)	(6)
	Chinese	Math	English	Chinese	Math	English
Proportion female peer	0.255	0.653**	0.571**	0.284	0.595**	0.552**
	(0.262)	(0.274)	(0.262)	(0.238)	(0.260)	(0.243)
Female student	0.590***	0.163***	0.564***	0.570***	0.133***	0.538***
	(0.024)	(0.028)	(0.024)	(0.024)	(0.028)	(0.025)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	No	No	Yes	Yes	Yes
Observations	8,665	8,664	8,663	8,246	8,245	8,244
R-squared	0.088	0.008	0.079	0.114	0.041	0.118

Notes: The dependent variable is the test score for each subject, normalized by school and grade. Individual control includes: student's age, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and repeated a grade in primary school. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Table 6. Mechanism: Classroom Environment

	(1)	(2)
Classmates are friendly	0.395*	0.430**
	(0.215)	(0.216)
Observations	8,814	8,402
R-squared	0.054	0.061
Classroom environment	0.809**	0.785**
	(0.369)	(0.373)
Observations	8,815	8,401
R-squared	0.117	0.122
School fixed effects	Yes	Yes
Grade fixed effects	Yes	Yes
Individual controls	No	Yes

Notes: Each cell represents a separate regression with specification (1), and the dependent variable is the outcome as listed above. Individual control includes: student's age, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and repeated a grade in primary school. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Table 7. Mechanism: Teacher-Student Interaction

	(1)	(2)
Head teacher: praises	0.522	0.435
	(0.334)	(0.335)
Observations	8,808	8,396
R-squared	0.095	0.102
Head teacher: criticisms	-0.171	-0.146
	(0.212)	(0.223)
Observations	8,815	8,402
R-squared	0.049	0.050
School fixed effects	Yes	Yes
Grade fixed effects	Yes	Yes
Individual controls	No	Yes
Subject teacher: asking questions	0.782***	0.718***
	(0.270)	(0.272)
Observations	26,472	25,254
R-squared	0.099	0.104
Subject teacher: praises	0.506	0.467
	(0.341)	(0.346)
Observations	26,467	25,243
R-squared	0.107	0.115
Subject fixed effects	Yes	Yes
School fixed effects	Yes	Yes
Grade fixed effects	Yes	Yes
Individual controls	No	Yes

Notes: Each cell represents a separate regression with specification (1), and the dependent variable is the outcome as listed above. Individual control includes: student's age, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and repeated a grade in primary school. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Table 8. Mechanism: Teacher's Job Satisfaction

	(1)	(2)
Satisfied with being a teacher	0.876	0.748
	(0.538)	(0.552)
Observations	459	439
R-squared	0.251	0.286
Feel school's atmosphere is good	0.241	0.269
	(0.530)	(0.522)
Observations	459	439
R-squared	0.541	0.571
Students in class respect me	0.773	0.700
	(0.482)	(0.514)
Observations	457	438
R-squared	0.430	0.454
Subject fixed effects	Yes	Yes
School fixed effects	Yes	Yes
Grade fixed effects	Yes	Yes
Teacher controls	No	Yes

Notes: Each cell represents a separate regression in which the dependent variable is the teacher level outcome as listed above, and the independent variable is the proportion of female students in class. Teacher control includes: subject teacher's age, teaching experience, education, and dummy variables indicating married, has certificate credential, and graduated from a normal college. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Table 9. Gender Peer Effect on Academic Outcomes by Student Gender

	Test score			Self-assessment score		
	(1)	(2)	(3)	(4)	(5)	(6)
	Female	Male	Difference	Female	Male	Difference
Proportion female peer	0.071	0.959***	-0.888**	-0.097	0.079	-0.176
	(0.233)	(0.296)	[0.001]	(0.196)	(0.217)	[0.509]
Subject fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,217	12,518		12,429	12,848	
R-squared	0.061	0.059		0.044	0.042	

Notes: Test score and self-assessment score are normalized by subject, grade and school, to obtain a mean of zero and standard deviation of one. Individual control includes: student's age, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and repeated a grade in primary school. In column (1), (2), (4) and (5), Standard errors are clustered at class level and reported in parentheses. In column (3) and column (6), p-values are reported in brackets. ***significant at the 1% level, **5% level, *10% level.

Table 10. Gender Peer Effect on Noncognitive Behavioral Outcomes by Student Gender

	Mental stress					Social acclimation and satisfaction					Disciplinary Problems		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: Female													
									Social activity with classmate				
	AES	Depressed	Blue	Unhappy	Pessimistic	AES	Fulfilling of life	Confident abt future	Public enrichment	Private recreation	AES	Late for school	Skip class
Proportion Female peer	-0.586**	-0.300	-0.593**	-0.748**	-0.711**	-0.161	0.180	0.291	0.777***	-0.161	-0.161	-0.330	0.032
	(0.272)	(0.295)	(0.291)	(0.288)	(0.339)	(0.201)	(0.192)	(0.260)	(0.274)	(0.201)	(0.201)	(0.244)	(0.178)
School fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observation	4,055	4,098	4,087	4,096	4,080	4,109	4,144	4,064	4,057	4,144	4,144	4,147	4,146
R-squared		0.078	0.073	0.083	0.059	0.083	0.108	0.239	0.254			0.088	0.071
Panel B: Male													
									Social activity with classmate				
	AES	Depressed	Blue	Unhappy	Pessimistic	AES	Fulfilling of life	Confident abt future	Public enrichment	Private recreation	AES	Late for school	Skip class
Proportion Female peer	-0.010	0.075	0.345	-0.252	-0.221	0.581***	0.699**	0.789***	0.200	0.621**	-0.559**	-0.330	-0.860***
	(0.239)	(0.296)	(0.279)	(0.284)	(0.264)	(0.135)	(0.272)	(0.222)	(0.314)	(0.270)	(0.219)	(0.245)	(0.276)
School fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observation	4,158	4,196	4,181	4,189	4,180	4,060	4,256	4,283	4,155	4,134	4,287	4,293	4,288
R-squared		0.044	0.054	0.043	0.043		0.097	0.085	0.213	0.215		0.095	0.098

Table 10. Gender Peer Effect on Noncognitive Behavioral Outcomes by Student Gender (Continued)

Panel C: Gender difference										
	Mental stress				Social acclimation and satisfaction				Disciplinary Problems	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Depressed	Blue	Unhappy	Pessimistic	Fulfilling of life	Confident abt future	Social activity with classmate: Public enrichment	Social activity with classmate: Private recreation	Late for school	Skip class
Gender Difference	-0.375	-0.938**	-0.496	-0.49	-0.426	-0.609**	0.091	0.156	0	0.892**
P-value	[0.315]	[0.015]	[0.193]	[0.217]	[0.174]	[0.027]	[0.744]	[0.601]	[0.999]	[0.012]

Notes : Individual control includes: student's age, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and repeated a grade in primary school. In panel A and Panel B, standard errors are clustered at class level and reported in parentheses. In Panel C, p-values are reported in brackets.***significant at the 1% level, **5% level, *10% level.

Appendix Tables

Table A1. Robustness Check: Nonlinear Effects of Gender Peer on Academic and Noncognitive Outcomes

	Academic outcomes		Mental stress				Social acclimation and satisfaction				Disciplinary Problems	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
									Social activity with classmate			
	Test score	Self-assessed score	Depressed	Blue	Unhappy	Pessimistic	School is fulfilling	Confident abt future	Public enrichment	Private recreation	Late for school	Skip class
20-40% female peer	0.302***	-0.067	-0.560***	-0.798***	-0.453**	-0.566	0.759***	0.191*	-0.189	-0.136	0.022	0.087
	(0.091)	(0.135)	(0.164)	(0.191)	(0.182)	(0.408)	(0.170)	(0.098)	(0.208)	(0.146)	(0.153)	(0.116)
40-60% female peer	0.246***	-0.089	-0.510***	-0.781***	-0.473**	-0.607	0.750***	0.254**	-0.153	-0.051	-0.001	-0.009
	(0.094)	(0.136)	(0.167)	(0.193)	(0.185)	(0.412)	(0.180)	(0.102)	(0.235)	(0.182)	(0.162)	(0.116)
60-80% female peer	0.369***	-0.068	-0.488***	-0.765***	-0.505***	-0.634	0.792***	0.362***	-0.134	0.016	-0.029	-0.037
	(0.113)	(0.147)	(0.176)	(0.201)	(0.191)	(0.415)	(0.183)	(0.106)	(0.235)	(0.189)	(0.166)	(0.115)
80-100% female peer	1.044***	0.513***	-0.825***	-0.391*	-0.010	-0.028	0.924***	0.294***	-0.756***	-0.511**	-0.544***	-0.138
	(0.119)	(0.149)	(0.180)	(0.206)	(0.193)	(0.417)	(0.187)	(0.109)	(0.241)	(0.197)	(0.172)	(0.119)
Female student	0.411***	0.174***	0.156***	-0.037	0.048**	-0.067***	0.083***	-0.026	0.045**	0.019	-0.087***	-0.095***
	(0.021)	(0.018)	(0.025)	(0.025)	(0.023)	(0.021)	(0.022)	(0.023)	(0.021)	(0.027)	(0.027)	(0.023)
Subject fixed	Yes	Yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
School fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24,735	25,277	8,294	8,268	8,285	8,260	8,365	8,427	8,219	8,191	8,440	8,434
R-squared	0.080	0.024	0.054	0.054	0.052	0.043	0.084	0.084	0.219	0.223	0.075	0.076

Notes: The omitted category is the share of female peer in classroom of 0-20%. Test score and self-assessment score are normalized by subject, grade and school, to obtain a mean of zero and standard deviation of one. Individual control includes: student's age, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and repeated a grade in primary school. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Table A2. Robustness Check: Sample Attrition

	Academic outcomes		Mental stress				Social acclimation and satisfaction				Disciplinary problems	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Test score	Self-assessment score	Depressed	Blue	Unhappy	Pessimistic	School is fulfilling	Confident abt future	Social activity with classmate: Public enrichment	Social activity with classmate: Private recreation	Late for school	Skip class
Proportion female peer	-0.013	-0.035	-0.017	-0.026	-0.015	-0.016	-0.059	-0.037	-0.002	-0.010	-0.031	-0.046
	(0.051)	(0.039)	(0.051)	(0.050)	(0.050)	(0.048)	(0.039)	(0.040)	(0.046)	(0.043)	(0.038)	(0.039)
Own gender control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Subject fixed effects	Yes	Yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26,730	26,730	8,910	8,910	8,910	8,910	8,910	8,910	8,910	8,910	8,910	8,910
R-squared	0.043	0.029	0.029	0.032	0.027	0.029	0.032	0.031	0.031	0.030	0.030	0.029

Notes: Each column represents a separate regression with specification (1), and the dependent variable is the attrition dummy for each outcome variable as listed above. Standard errors are clustered at class level and reported in parentheses. ***significant at the 1% level, **5% level, *10% level.