# The Effect of Teacher Gender on Students' Academic and Noncognitive Outcomes

Jie Gong, Yi Lu, and Hong Song

National University of Singapore, Singapore*

This Version: September 2016

**Abstract**

This paper examines the role of teacher gender in education production. We extend student outcomes from traditionally focused academic achievement to noncognitive outcomes. Drawing on a representative survey of middle-school students and teachers in China, we focus on schools in which student-teacher assignments are random. Our results show that having a female teacher raises girls' test scores and improves both their mental status and social acclimation relative to boys. We further find that the effect on test scores is larger for students whose parents are less educated or have migrated out. There is evidence that female teachers provide feedback differently to girls and boys, and that having a female teacher alters girls' beliefs about commonly held gender stereotypes and increases their motivation to learn.

**Keywords**: Gender difference; teacher gender; noncognitive skills

**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 Link, National University of Singapore, Singapore 117570.

# 1 Introduction

The past two decades have witnessed the reversal of a remarkable gender gap in education. In the US and many other developed countires, girls outperform boys in reading scores, top GPA distribution in high school and college attendance (Campbell 2000; Vincent-Lancrin 2008; Fortin, Oreopoulos and Phipps 2015). The international phenomenon of female dominance in academic achievements raises new questions about the causes of this disparity. Are the differences primarily biological or nurtured by social influences? Are certain educational inputs, either from parents or teachers, more suitable for girls than boys? For instance, is it only by coincidence that females comprise the majority of teachers in elementary and secondary schools during the period when girls achieve more and behave better than boys?

To gain insight into this question, our paper examines whether teacher gender impacts the learning of girls and boys at the middle-school level. A growing number of studies have attempted to document the effect of teacher gender at the primary-school (e.g., Winters, Haight, Swaim and Pickering 2013; Antecol, Eren and Ozbeklik 2015); secondary-school (e.g., Ehrenberg, Goldhaber and Brewer 1995; Nixon and Robinson 1999; Dee 2005, 2007; Winters, Haight, Swaim and Pickering 2013); and college level (e.g., Canes and Rosen 1995; Rothstein 1995; Neumark and Gardecki 1998; Bettinger and Long 2005; Hoffmann and Oreopoulos 2009; Carrell, Page and West 2010). These studies tend to focus on the impact on student academic performance, and find that having a female teacher improves female students' educational outcomes. In contrast, relatively less research has examined the effects of teacher gender on students' noncognitive outcomes—e.g., attitudes, preferences, and socio-emotional factors. The mechanisms behind teacher-student gender interactions have also limited empirical support.

We attempt to fill these gaps in the literature. First, we extend student outcomes from traditionally focused academic achievements to noncognitive outcomes. Since Jencks et al. (1979), a number of 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; Flossmann, Piatek and Wichert 2008; Segal 2013; Bertrand and Pan 2013). To understand the role of school influence on noncognitive outcomes, we focus on the middle-school period, as this is the age at which noncognitive skills are thought to develop and mature (Borghans, Duckworth, Heckman and Ter Weel 2008; Heckman and Kautz 2013).

Our second contribution is to uncover the mechanisms by which gender interactions between teacher and student lead to differential outcomes for girls and boys. Do teachers have a preference for same-gender students so that they provide differential attention or

responses to students depending on gender? Or, given the same teaching behaviors, do girls and boys react and respond to instruction in different ways? Understanding these mechanisms is important when designing policies to mitigate gender differences at school (Dee 2004, 2007; Carrell, Page and West 2010). It may be especially relevant when it is difficult to change the gender composition mix of a school's instructors. Our research provides insight into this question by using unique data on teacher behavior and student beliefs to directly test how a teacher's behavior differs by student gender as well as how students' beliefs and motivations related to learning are influenced by their teacher's gender.

We use a nationally representative survey of middle-school teachers and students in China and focus on schools in which the assignment of students to classrooms is random. This allows us to mitigate potential selection problems—e.g., high-achieving or better-motivated female students are more likely to be assigned to female teachers. By comparing the outcomes of girls and boys taught by female versus male teachers, we are able to estimate whether, and how, teacher gender affects female and male students' outcomes differently.

Our results show that although teacher gender has little impact on boys, it has a strong effect on girls, in both academic and noncognitive outcomes. On academic performance, our estimates confirm a reverse gender gap: Girls outperform boys on test scores. What is immediately remarkable is how female teachers increase the gender gap by improving girls' scores relative to boys'. Teacher gender has little impact on boys' test scores, but girls' scores increase by about 14.4% of a standard deviation when taught by a female. The results are robust to the inclusion of various student- and teacher-level controls.

Notably, having a female teacher also improves girls' self assessment of their learning. When taught by a male teacher, although girls' absolute grades are higher than boys', their self-assessed outcomes are not, which reflects a gender gap in confidence about academic achievement. Having a female teacher significantly increases girls' self-assessed scores relative to boys', and the magnitude of the difference is even larger than the effect on girls' actual test scores.

The pattern for students' noncognitive outcomes is subtler. If taught by a male, girls are more likely to feel depressed, blue, or unhappy at school than boys. Female teachers can overturn the gender gap: They do not seem to affect boys' outcomes, but significantly improve girls' mental status. In addition, we find strong evidence that female teachers improve girls' overall satisfaction at school, as well as their social acclimation with classmates.

Examining the mechanisms that drive these results, we find evidence for both teacher- and student-based channels. Specifically, we find that female teachers tend to ask more questions, give more praise, and make fewer critical comments to girls. In addition, we find that the presence of a female teacher counters the perception that girls are not as strong in

math, and better motivates girls to study the subject. Overall, our findings support both the stereotype-threat hypothesis and role-model theories.

In the last part of our empirical analysis, we examine whether students from lower socioeconomic backgrounds are more sensitive to the influence of teacher gender. Our findings suggest that when the mother is less educated, the student is an ethnic minority, or the parent migrated out and left the student behind, teacher gender has a much stronger impact on students' test scores. These patterns are consistent with the premise that disadvantaged students receive less parental investment, so school-based influences play a more important role in their skill development.

Our findings provide new evidence that teacher gender has a significant influence on student outcomes at the middle-school level. This reinforces previous findings regarding the impact of teacher gender on student grades, and also contributes to the literature on the effect on middle school students in particular. Some studies have found that having a female teacher positively affects female middle-school students' achievement (Nixon and Robinson 1999; Dee 2007; Winters, Haight, Swaim and Pickering 2013), while others have found no effect (Ehrenberg, Goldhaber and Brewer 1995). Our study also extends previous research by using a setting with randomized student assignment. To our knowledge, only Dee (2004), Carrell, Page and West (2010), and Antecol, Eren and Ozbeklik (2015) use randomized or experimental data to account for endogenous assignment, using data from the Project STAR, the US Air Force Academy, and the National Evaluation of Teach for America (NETFA), respectively.[1] Our estimates are based on randomized student assignments and a more representative sample of middle-school students and teachers, and therefore are arguably more general from a policy perspective.

More broadly, we also contribute to the literature on how school environment influences noncognitive skills. Previous studies highlight the central role of school choice and school social networks. For instance, Cullen, Jacob and Levitt (2006) find that gaining access to sought-after public schools improves students' behavioral outcomes, as evidenced by a lower level of disciplinary incidents and arrest rates. Angrist, Bettinger and Kremer (2006) find that winners of vouchers for private schools in Colombia were more likely to finish the 8th grade, worked less, and were less likely to marry or cohabit as teenagers than those who did not win vouchers. Lavy (2010) evaluates a program of free choice among public schools in Tel Aviv, Israel, and finds better student behavioral outcomes, such as more positive teacher-student relationships, better social acclimation and satisfaction at school, and less

[1] Several studies deal with the selection problem by making within-student and within-teacher comparisons. See, for example, Dee (2007), Hoffmann and Oreopoulos (2009) and Fairlie, Hoffmann and Oreopoulos (2014).

violence and disruption in the classroom. Lavy and Sand (2015) exploit conditional random assignment in middle schools in Tel Aviv, Israel, and find that social networks have positive effects on students' noncognitive behavioral outcomes: greater cooperation, reduction in violent behavior and improvements in social satisfaction in class. Our study complements the literature by providing evidence of how one component of the educational experience, teacher gender, influences student noncognitive outcomes. An important advantage of our study is that our findings are not limited to a particular type of school—i.e., public or private—because the estimates stem from within-school comparisons.

# 2 Data and Variables

In the Chinese educational system, middle-school students are assigned to classrooms at the beginning of the 7th grade and take the same courses throughout their three years of middle school. Students are required to take three core subjects—Chinese, mathematics, and English—and a set of subsidiary subjects. During a regular school day, students remain in the same classroom all day and different teachers come to the classroom to deliver subject-specific lectures. A head teacher, who is usually one of the core-subject teachers, oversees the activities and individual student progress for a given classroom. While the subject teachers ensure that students achieve their learning goals, the head teacher is also responsible for all class matters and students' social lives—e.g., setting seating plans, organizing extracurricular events, and overseeing student discipline. In addition, the head teacher regularly gives feedback to students and their parents regarding academic performance and behavior.

We obtained data from the 2014 China Education Panel Survey (CEPS), a nationally representative survey that covers middle schools from 28 counties and city districts.[2] Our sample includes 8,988 students in the 7th and 9th grades, across 208 distinct classrooms and 67 schools. Data for each student's demographic characteristics are collected from student questionnaires. Table 1 reports the summary statistics for our main outcome and control variables.

[Insert Table 1 here]

Our academic performance measures are student exam scores in the core subjects, provided by their respective school administration offices. These subjects—Chinese, mathematics, and English—are the main components in the standards tests for admission to senior

[2] CEPS is the first and largest nationally representative education survey in China. Officially started in 2013, the survey applies a stratified sampling design in which four middle schools and four classrooms within each school are chosen to represent a given county or city district.

high school. Within a school, teachers who teach the same subject use the same syllabus and give the same exams during a common testing period. Test scores in the core subjects are therefore a consistent measure of academic achievement across students from the same grade in the same school.[3] We supplement our test score measures with self-assessed performance scores collected from the student questionnaire, on which students were asked to report whether they have difficulty in learning each subject on a scale from 1 (*a lot*) to 4 (*not at all*).

To measure student noncognitive outcomes, we use student responses to eight questionnaire items. Four questions relate to feelings of stress; specifically, whether during the previous seven days the student felt (1) depressed, (2) blue, (3) unhappy, or (4) that life is meaningless. The remaining questions relate to a student's perceptions of his/her social acclimation and general satisfaction levels at school; specifically, (5) whether the student feels that school life is rich and fulfilling, (6) whether the student feels confident about his/her future, (7) how often the student goes to museums, zoos, or science parks with classmates from school, and (8) how often the student goes to movies, plays, or sporting events with classmates from school.[4]

Our main teacher-level variable of interest is gender. In addition, we obtain from teacher questionnaires information about which subject each respondent taught; his/her education level, marital status, and years of teaching experience at the current school; and years in the profession. Our summary statistics show that 64.9% of head teachers and 77.9% of subject teachers in our sample are female. Within each core subject, female teachers are also more common than male teachers: 75.9% in Chinese, 65.9% in math, and 91.6% in English.[5]

# 3 Estimation Strategy

## 3.1 Class Assignment and Regression Sample

Our research question concerns the effect of teacher gender on student outcomes. Understanding how students are matched to teachers and classrooms is therefore critical to our

[3]Exams are graded in a rigorous and consistent manner with each student's name, class, and ID information hidden from the grader. Within a grade at the same school, teachers divide the grading work so that the same question is typically graded by the same teacher using a consistent standard.

[4]For items (1) to (4), respondents were asked to report the frequency of incidents on a scale from 1 (*never*) to 5 (*always*). For items (5) and (6), 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*). In regression analysis, the variables are normalized to have means of zero and standard deviations of one.

[5]Our observed female dominance in secondary school teachers is a global phenomenon. See, for example, Holmlund and Sund (2008) and Winters, Haight, Swaim and Pickering (2013).

estimation and analysis. Middle schools in China use various methods to assign students. In some schools, prior to beginning their first academic year, students take placement exams and their scores are used to assign them to classrooms. In other schools, students are assigned based on local residency.

More recently, an increasingly large number of schools have begun to employ random assignment to place students to classrooms. This approach is heavily promoted by the Ministry of Education as ensuring equal and fair opportunity for all students during their compulsory education years (through the 9th grade). Schools that use randomized assignments typically rely on a computer program that can incorporate information on class size, gender, migrant status, and other dimensions to ensure proper balance in the randomization process. Alternatively, if enrollment is small and manageable, parents of incoming students are invited to draw lots to determine their child's class placement. In these cases, once student assignments have been determined, the head and subject teachers also draw lots to determine which classrooms they will teach and manage.

In this study, we focus on schools in which students are randomly assigned to classrooms. We require that a school's class assignment procedure to meet three conditions: (i) the school principal reports that students are randomly assigned to classrooms; (ii) after students have been assigned to classrooms at the beginning of the 7th grade, the school does not rearrange their classes for grades 8 and 9; and (iii) all head teachers in the same grade report that students in the respective grade are *not* assigned by test scores.[6] Using our criteria, we find that 59.8% of the schools in the 2014 CEPS database assign classes randomly, translating into a sample of 8,988 students across 208 classrooms and 67 schools. Since each student in our sample is randomly matched to both a head teacher and subject teachers, and stays in the same class for the next three years, our sample mitigates any potential concerns regarding self-selection of students to classrooms or teachers.

To verify the randomness of class assignment for our sample, we conduct a balancing test between classrooms with a female head teacher and those with a male head teacher using several baseline characteristics (i.e., values that are determined before the class assignment). If the assignment process is truly random, these two groups of students should be similar across the observed characteristics. We include the following student characteristic variables: *Female* (1 if a student is a female and 0 otherwise), *Age*, *Minority* (1 if a student belongs

[6]The criteria are based on reports in the teacher questionnaire. For the first condition, all school principals were asked to report which of the following assignment rules they used 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 (c). Second, the same principals were asked whether their schools rearrange classes for grades 8 and 9; we exclude those that do so. Finally, each head teacher was asked whether students in the grade level taught are assigned by test scores; again, we drop the entire grade if any head teacher answers "yes".

to a minority ethnic group and 0 otherwise), *Local Residence* (1 if a student is local 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 primary school and 0 otherwise), *Academic Ranking in Primary School* (ranking of academic performance in primary school), *Mother Education* (mother's years of schooling), and *Father Education* (father's years of schooling). Our classroom-related baseline variables are: *Proportion of Other Female Subject Teachers* (percentage of female subject non-head teachers), *Chinese Teacher's Teaching Experience* (years of teaching), *Math Teacher's Teaching Experience* (years of teaching), *English Teacher's Teaching Experience* (years of teaching), and *Class Size* (number of students).

Panel A of Table 2 reports test results using student characteristic variables. For each variable, we report the mean and standard deviation of classes with female head teachers and those with male head teachers in columns 1 and 2, respectively. Column 3 presents raw differences between the two groups. As random assignment is performed within a school, we further report conditional differences controlling for school fixed effects in column 4, which allows us to control for all possible selection by students among schools in the cross-sectional data. While there are several statistically significant unconditional differences across the two groups of classes, these differences disappear once we control for school fixed effects. Also, all of the differences are economically insignificant. These results suggest that the student characteristics in our sample are well balanced across classes managed by female versus male head teachers.

[Insert Table 2 here]

Table 2, Panel B, shows the balance of classroom characteristics. When we include school fixed effects, none of the five classroom variables—proportion of female subject teachers, experience of the three core-subject teachers, or class size—shows any statistically or economically significant differences between classes with female head teachers and male head teachers.

As the focus of our analysis is on the interactions between student gender and teacher gender, there may still be concern that the teacher-student match is based on unobservable along both teacher and student gender. For example, if within the same grade or school the principal puts strong girls in a class with a female teacher but also assign the weak boys to her class, we could still get seemingly balanced results on female teacher in Table 2 as the selection cancels out when one pools both student genders. To this end, we further check the balance of gender difference between classes with female head teachers and those with male head teachers in Appendix Table A1. Specifically, we regress all of the students'

pre-determined characteristics used in Table 2 on student gender, teacher gender, and their interactions, along with school fixed effects. This is the same estimation specification we employ to estimate teacher gender effects (see the next section for more details). With the exception of ethnicity, all of coefficients on female student and female teacher interactions are insignificant. In addition, all magnitudes are economically insignficant.

Taken together, these results largely confirm that students taught by a female head teacher have characteristics and classroom environments similar to those taught by a male head teacher. Nonetheless, we include all these student-level predetermined characteristics in the later analyses to further improve the balance between the two groups of classrooms, as well as our estimation efficiency.

## 3.2 Estimation Framework

To estimate the effect of teacher gender on student academic and noncognitive outcomes, we use the following regression model.

$$Y_{ics} = \alpha + \beta_1 Femstud_{ics} + \beta_2 Femteach_{cs} + \beta_3 Femteach_{cs} \times Femstud_{ics} + W'_{ics}\phi + D_s + \varepsilon_{ics} \quad (1)$$

where $Y_{ics}$ is the outcome measure for student $i$ in class $c$ of school $s$; $Femstud_{ics}$ is an indicator for whether the student is female; $Femteach_{cs}$ is an indicator for whether the teacher for class $c$ of school $s$ is female; and $\varepsilon_{ics}$ is the error term. We cluster standard errors at the class level, accounting for any correlation in outcomes for students in the same class.

The $\beta$ coefficients are the main coefficients of interest. $\beta_1$ captures the difference in mean academic performance and noncognitive outcomes between female and male students when they have a male teacher. $\beta_2$ captures the value added by having a female teacher, and $\beta_3$ captures the effect of teacher gender on student gender differences. An unbiased estimation of the $\beta$ coefficients requires that conditional on the controls, $Femteach_{cs}$ and $Femstud_{ics}$ are uncorrelated with the error term $\varepsilon_{ics}$. These identifying assumptions hold in a setting of random class assignment, which is verified in the previous section. In particular, as random class assignment is conducted within a school and students may nonrandomly select their middle school, we include school fixed effects $D_s$ in the regressions to control for all school-level factors in our cross-sectional data that may influence students' school selection decisions.

# 4 Main Results

## 4.1 Teacher Gender and Student Academic Performance

Table 3 presents our estimates of the effects of core-subject teachers' gender on student academic performance. Columns 1 and 2 present results using test scores as the performance measure while columns 3 and 4 present results using self-assessed performance as the outcome variable. Both test scores and self-assessed scores are normalized by subject, grade and school; therefore, both variables have means of zero and standard deviations of one. We include subject fixed effects and school fixed effects in all regressions, and student-level control variables for the results in columns 2 and 4.

[Insert Table 3 here]

Baseline estimates on test scores (columns 1 and 2) establish two facts: Girls outperform boys in course tests, and the gender difference is enhanced by teacher gender. The coefficients on female-student dummy are positive and significant, which confirms that girls achieve higher scores than boys (19.4% of a standard deviation). This is consistent with previous evidence on female dominance in academic performance.[7]

The more intriguing findings are that female teachers enlarge the gender difference—that is, they improve girls' scores more relative to boys. The estimated coefficients on the female-teacher dummy variable are negative but insignificant, suggesting that teacher gender has little impact on boys' test scores. In contrast, estimated coefficients on the interaction between female teacher and female student are positive—14.4% of a standard deviation—and precisely estimated. These results imply that female teachers substantially improve girls' test scores relative to boys. The absolute gain to girls from having a female teacher is about 10.4% of a standard deviation (0.144-0.040).

An interesting finding from students' self-assessment scores (columns 3 and 4) is that although girls' absolute grades are higher than boys', their self-assessed outcomes are not. This reflects a gender gap in confidence about academic achievement. Female teachers help reduce the gap: They raise girls' self-assessment by 27.3% of a standard deviation relative to boys, which translate into an absolute gain to girls of about 15.6% of a standard deviation (0.273-0.117). The presence of female teachers seems to have a greater effect on girls' self-

[7] A similar reverse gender gap has been found among students in the U.S. and Sweden. US educational statistics report that between 1973 and 1999, female students consistently dominated male students in reading scores, while the gender gap in mathematics decreased and became insignificant (Campbell 2000). The Swedish National Agency for Education (2004) reports that girls score significantly higher than boys on Swedish and English tests, and there is no obvious gender difference in mathematics scores (Holmlund and Sund 2008).

assessment than it does on actual scores, suggesting that teacher gender plays an important role in students' mental and behavioral outcomes.

## 4.2 Teacher Gender and Students' Noncognitive Outcomes

To estimate the effects of the head teacher's gender on students' noncognitive outcomes, we normalize each of the eight noncognitive outcome variables to obtain a mean of zero and standard deviation of one.[8] When reporting the results, we group the outcome variables into two sets: the level of mental stress and the level of social acclimation and satisfaction. Following Kling, Liebman and Katz (2007), we also calculate the average effect size (AES) for each category to supplement the estimated effect on a specific outcome measure.[9]

Table 4A presents the estimated effects on students' mental stress. First, the estimated coefficients on the female student dummy variables are mostly positive with the exception of a negative yet insignificant coefficient for the likelihood of feeling pessimistic. These results indicate that girls are more likely to feel depressed, blue, or unhappy than boys when their head teacher is male. Second, estimates of the female-teacher dummy variable are negative, suggesting that a female head teacher tends to improve boys' mental status, though this is not statistically significant. Finally, and most importantly, estimated coefficients on the female teacher $\times$ female student interaction are negative, indicating that having a female head teacher significantly reduces girls' mental stress relative to boys. Estimates are statistically significant for students' likelihood of feeling depressed, blue, or unhappy, as is the AES.

[Insert Table 4A here]

Table 4B shows the estimated effects on students' social acclimation and satisfaction. When controlling for student-level characteristics, estimated coefficients on the female student dummies are negative for most measures (except for "fulfilling of life" which is close to zero and not statistically significant), suggesting that girls are less satisfied with school life, less confident about the future, and less likely to hang out with classmates when their head teacher is male. In contrast, having a female teacher is associated with greater satisfaction,

[8] We focus on the gender of a student's head teacher as the head teacher is more involved in students' activities at school than are subject teachers. Also, a student's noncognitive outcomes do not vary by subject. Nevertheless, as a robustness check, we include the proportion of female teachers (both head and subject teachers) as an alternative measure of teacher gender and report the results in Section 4.3 and Appendix Table A3.

[9] Specifically, we define the average effect size (AES) of the interaction between female teacher dummy and female student dummy on category $c$ as $AES = \frac{1}{n_c}\sum_{n=1}^{nc} e_{kc}/\sigma_{kc}$, where $n_c$ is the number of outcomes in category $c$ (for our noncognitive categories, 4), $e_{kc}$ is the estimated effect for outcome $k$ of category $c$, and $\sigma_{kc}$ is the standard deviation of the outcome variables.

confidence, and acclimation for boys, although these effects are not significant. Finally, we find that our estimates for female teacher × female student interaction are positive and significant, suggesting that relative to boys, girls have greater satisfaction and acclimation at school when their teacher is female.

[Insert Table 4B here]

One concern regarding students' noncognitive outcomes is that they might be correlated with students' academic performance. For instance, high achievers are more likely to feel confident and satisfied in school, whereas students who perform worse on academic tests are more prone to mental stress. If such correlations are strong, then a teacher's gender influences students' noncognitive outcomes primarily through the effects on academic performance. To separate the direct effect from the indirect effect via academic performance, we add students' test scores as controls for noncognitive outcomes. As shown in Table 4C, the estimates remain almost the same after controlling for test scores. In other words, holding test scores constant, having a female teacher still contributes to positive feelings of mental health, school satisfaction, and social acclimation for girls relative to boys.

[Insert Table 4C here]

The findings across various specifications suggest that girls have lower noncognitive outcomes than boys—worse mental status, weaker confidence and lower satisfication at school—when they are assigned to male teachers. A female teacher can reduce, or even overturn the gender gap by improving girls' noncognitive measures relative to boys.

## 4.3 Robustness Checks

We conduct several robustness checks on various estimation concerns, including an alternative measure of teacher gender, adding more fixed effects in our specification, and testing sample attrition.

First, in our main analyses, we use the gender of the core subject teacher in estimating the effect of teacher gender on student academic performance. Considering the intensive interaction between head teacher and students, it is also possible that the gender of a student's head teacher affects that student's academic achievement. To address this possibility, we include a dummy variable indicating the gender of the head teacher as well as its interaction with student gender. The results in Appendix Table A2 show that estimates of the subject-teacher gender effect (coefficients of the interaction between female subject teacher and female student) remains significant and are even slightly larger than our baseline esti-

mate reported in Table 3. The coefficients on female head teacher are positive and significant, but those on the female head teacher and female student interaction term are negative and insignificant. Thus, we conclude that while the gender of a student's head teacher affects students' academic achievement, it does not seem to enlarge or correct the gender gap.

Second, we use the proportion of female teachers in a classroom as an alternative measure of teacher gender. Our baseline specification uses a dummy variable to indicate whether the head or subject teacher is female. Here, we calculate the proportion of female teachers by teacher gender and self-reported working hours with students. The more hours a female teacher spends with students, the more intensive the female teacher-student interaction. Appendix Table A3 reports the results. Consistent with baseline results, we find that the larger share of student-teacher interaction spent with female teachers, the less likely that girls feel depressed, blue, or unhappy and the more likely that girls have better acclimation and satisfaction relative to boys. Overall, we find that our results are not sensitive to the alternative measure of teacher gender.

For our next set of robustness checks, we test whether our main findings are robust to the inclusion of (1) student fixed effects; (2) grade (7th or 9th) fixed effects; and (3) teacher-level control variables.[10] As shown in Appendix Tables A4-A5, the estimated effect of teacher gender remains significant and similar to the baseline estimates.

Last, we address the potential issue of sample attrition; that is, missing values on student and teacher variables. Specifically, our estimates would be biased if teacher gender affects the likelihood of missing values. We regress an attrition dummy on teacher's gender, student gender, and school fixed effects. As shown in Appendix Table A6, the coefficients on teacher gender are close to zero and not statistically significant, indicating that our results are unlikely to be driven by teacher gender.

# 5 Mechanism

Our main results in Tables 3 and 4 suggest that relative to male students, female students' test scores and noncognitive outcomes are substantially better when they are taught by female teachers. In this section, we investigate possible mechanisms through which teacher gender affects female and male students outcomes differently.

[10] We include student fixed effects in estimating the impact on students' academic outcomes but not noncognitive skills. This is due to the fact that we observe a student's academic outcomes in multiple subjects, but there is no within-student variation in noncognitive measures.

## 5.1 Teacher Gender or Other Characteristics

We start by examining whether our documented gender differences in student performance are primarily driven by teacher gender *per se*, or, by other teacher characteristics correlated with teacher gender. For instance, students may respond differently to younger versus older teachers. To address the relevance of teachers' professional characteristics, we follow the approach used by Carrell, Page and West (2010).[11] Specifically, we add the interactions between student/teacher gender dummies with the following teacher-level variables: age, marital status, years of schooling, whether they earned their degree from a normal college, experience in the profession, experience as a head teacher, prior experience as a subject teacher, and credential status. Including these interactions can identify whether certain teacher characteristics affect female and male students differently. As shown in Table 5, we see that the estimate for teacher gender effect is similar to our baseline estimates. These results suggest that observed effects are driven primarily by a teacher's gender, rather than other teacher characteristics.

[Insert Table 5 here]

## 5.2 The Role of Teacher Gender: Teacher Behavior and Student Response

Given that teacher gender drives our results, we examine whether this effect is due to teacher behavior or student responses. Regarding teacher behavior, it is possible that teachers may prefer students of a particular race, ethnicity, or gender, leading to different behaviors depending on student characteristics. For instance, teachers may provide more feedback or time to students of the same gender (Jones and Dindia 2004).[12] Relatedly, female and male teachers may differ in their teaching styles or communication strategies, which may be better suited to students of their own gender.

Regarding student response, it is possible that students may perceive teachers differently based on the teacher's gender. For example, the role-model effect suggests that students feel more comfortable, inspired, and focused in class when a same-gender teacher is present, independent of the teacher's other qualities and behaviors.[13] Similarly, girls may feel less

[11] Carrell, Page and West (2010) also analyze a teacher's value added to distinguish teacher gender effects from other teacher characteristics effect. However, this approach requires multiple observations of a teacher, which is not applicable in our setting.

[12] In a similar view, Ferguson (1995) shows that in allocating class time and materials, teachers are more oriented to students of the same race.

[13] The literature on role-model effects is relatively rich; see, for example, Rothstein (1995), Canes and Rosen (1995), Jacobs (1996), and Diprete and Buchmann (2006).

threatened by stereotypes if they have a female teacher. In this case, a same-gender teacher will be perceived by students as a counterargument to commonly-held gender stereotypes—e.g., that girls are less talented than boys in mathematics—that may otherwise hinder girls in their ability to perform at their maximum level.[14]

Our rich data allow us to directly test and separate these two classes of explanations. We first examine whether teachers behave differently toward female and male students. On questionnaires, students were asked to describe their interactions with their subject and head teachers on dimensions such as class questioning and the provision of praise and criticism.[15] As shown in Table 6, we find that compared to a male teacher, a female head teacher is less likely to criticize girls than boys, and that a female subject teacher is more likely to question and praise girls than boys. The results remain similar after controlling for students' test scores. In other words, holding students' academic performance constant, there is still differential treatment by teachers based on student gender, in that female teachers provide more attention and positive feedback to girls than boys.

[Insert Table 6 here]

Next, we test students' perceptions of and responses to their teachers' gender. Table 7 presents evidence on the stereotype-threat hypothesis and role-model effects. Panel A reports the teacher gender effect on students' belief that "boys are more talented in learning math than girls." Our results show that when mathematics is taught by a female teacher, female students are less likely to agree with the stereotype about a gender gap in learning math. Exposure to a female teacher who counters the stereotype appears to increase female students' confidence, which may contribute to our observed improvements in performance.

To examine the relevance of role-model effects, we check whether a teacher's gender influences students' views about the usefulness of the subject for their future.[16] The role-model effect implies that students feel more inspired when taught by a same-gender teacher. From Table 7, Panel B, we see that while male students' perceptions do not vary significantly by teacher gender, female students are more likely to consider a subject relevant and useful

[14]The stereotype-threat hypothesis (Steele 1997; Spencer, Steele and Quinn 1999) is supported by experimental evidence (Steele and Aronson 1995), in which female students under-perform on a math test only when they are told that the test reveals gender differences.

[15]The four items ask students to rate how much they agree with the following statements on a scale from 1 (*strongly disagree*) to 4 (*strongly agree*): (1) The subject teacher always asks the student to answer questions in class, (2) The subject teacher always praises the student in class, (3) The head teacher always criticizes the student, and (4) The head teacher always praises the student. In the regression analyses, we normalize each variable to have a mean of zero and standard deviation of one.

[16]The survey item asks the students to rate, on a scale from 1 (*strongly disagree*) to 4 (*strongly agree*), how much the student agrees with the statement that "the subject is useful for my future."

for their future when the subject is taught by a female. We interpret this result as evidence to support the idea that a female teacher functions as a role model who inspires her female students.

[Insert Table 7 here]

Taken together, we find evidence for both teacher- and student-based channels. Our findings reveal teachers' differential behaviors depending on student gender; and their roles in countering stereotypes and inspiring same-gender students. Results are all robust to the inclusion of student characteristics and test scores. These findings, therefore, help to explain how female teachers improve girls' academic and noncognitive development.

# 6 Heterogeneous Effects

Our main findings, which are reported in Tables 3 and 4 capture the average effects of teacher gender on student outcomes. In this section, we examine whether the effects vary according to student background. In the spirit of Cunha and Heckman (2007), students develop skills through various investments—e.g., parenting, schooling, environment—and the returns from these investments are interdependent. In this case, a teacher's influence may vary by the student's parental investment.

We explore differential effects across three dimensions: parents' education (proxied by mother's years of schooling), student ethnicity, and whether the student's parents migrated out.[17] As shown in Table 8, we find that the female teacher effect is stronger for disadvantaged students. For instance, we find that having a female teacher raises left-behind girls' test scores by 29.6% of a standard deviation relative to boys, compared to the increase of 9.8% of a standard deviation observed for girls with both parents around. We further find that the coefficients on the "female teacher $\times$ female student" interaction are statistically different between the three pairs of subgroups: students whose mothers have more than 9 years of school versus those with less educated mothers, ethnic majority versus minority, and students who live with their parents versus those who were left behind by migrant parents.

[Insert Table 8 here]

In terms of noncognitive outcomes, we find that the teacher gender effects (coefficients of the interaction term) are higher for disadvantaged students across all the measures of

[17] For mother's education, we use 9 years of schooling as the cutoff since it is the sample median. Results are similar if we use father's education instead.

mental health (Table 9, Panel A), but comparisons of coefficients between subgroups are not statistically significant. The differences between subgroups in social acclimation and satisfaction at school are less pronounced (Table 9, Panel B). Because the effects do not differ statistically between subgroups, we should interpret heterogeneous effects with caution.

[Insert Table 9 here]

Overall, we find evidence that the female teacher effect on test scores is stronger for students from lower socioeconomic backgrounds, which may reflect less parental input for these students. As a result, these students are more sensitive to their school environments and more influenced by their teachers in learning and development outcomes. These results reinforce other studies' finding that disadvantaged students are more likely to be affected by the gender composition of their peers and class size (Angrist and Lavy 1999; Lavy and Schlosser 2011). An implication of the heterogenous effect is that school-related policies that target disadvantaged students may have large returns.

# 7 Conclusion

This paper sheds light on how teacher gender impacts student outcomes. We use random assignment of students to waive concerns about self-selection in student-teacher matching. Our data allow us to extend student outcomes from traditionally focused academic achievements to noncognitive and behavioral outcomes. A further advantage of our study is that various measures of teacher and student perceptions and behaviors are used to separate the possible mechanisms that drive the main effect of teacher gender.

Our findings suggest that having a female teacher improves both academic performance and noncognitive outcomes more for girls than for boys. On student test scores, our results show a reverse gender gap, which is enforced when taught by female teachers. This teacher gender effect is stronger among students from lower socioeconomic backgrounds. On noncognitive outcomes, girls tend to be more mentally stressed and less satisfied at school if they are taught by male teachers. However, female teachers can overturn this gender gap. Examining the mechanisms that drive these results, we find evidence of teacher behavior that varies depending on student gender, as well as student perception of teachers as role models and counterexamples to gender stereotypes.

Our study has a number of implications for educators and policymakers. First, our findings provide useful information for policy makers who seek to balance gender representation among middle-school teachers. Second, our results provide insight for teachers regarding the

differential impact of classroom actions such as questioning and praising or criticizing girls versus boys.

Our findings can also be used to broaden our understanding of the nature and sources of boys' academic difficulties (e.g., Bertrand and Pan 2013; Cornwell, Mustard and Van Parys 2013; Fortin, Oreopoulos and Phipps 2015). It is commonly observed and recognized that boys are more likely to have behavioral and socio-emotional difficulties such as problems, with self-control, which may limit their ability to obtain high grades. Our findings may help schools understand how teacher behaviors can contribute to a more gender-neutral school environment.

While this study has provided insight into the role of teacher gender at the middle school level, future research could investigate several important questions. For example, it would be valuable to study the long-term effects of teacher gender on students. Do these effects decline or persist to later education, and even adulthood? It would also be interesting to examine whether students are more affected by their most recent teachers or by teachers they had during pivotal stages of development, such as middle school. Finally, future studies could examine how teacher gender interacts with other school environment factors, such as peer influence. Answers to these questions would offer valuable insights into the effects of school context on children's development and long-term achievements.

# References

**Angrist, Joshua D and Victor Lavy**, "Using Maimonides' Rule to Estimate the Effect of Class Size on Student Achievement," *The Quarterly Journal of Economics*, 1999, *114* (2), 533–575.

**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.

**Antecol, Heather, Ozkan Eren, and Serkan Ozbeklik**, "The Effect of Teacher Gender on Student Achievement in Primary School," *Journal of Labor Economics*, 2015, *33* (1), 63–89.

**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.

**Bettinger, Eric P and Bridget Terry Long**, "Do faculty serve as role models? The impact of instructor gender on female students," *The American Economic Review*, 2005, *95* (2), 152–157.

**Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel**, "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 2008, *43* (4), 972–1059.

**Campbell, Jay R**, *Trends in Academic Progress: Three Decades of Student Performance*, US Department of Education, Office of Educational Research and Improvement, 2000.

**Canes, Brandice J and Harvey S Rosen**, "Following in Her Footsteps? Faculty Gender Composition and Women's Choices of College Majors," *Industrial & Labor Relations Review*, 1995, *48* (3), 486–504.

**Carrell, Scott E, Marianne E Page, and James E West**, "Sex and Science: How Professor Gender Perpetuates the Gender Gap," *The Quarterly Journal of Economics*, 2010, *125* (3), 1101–1144.

**Cornwell, Christopher, David B Mustard, and Jessica Van Parys**, "Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments: Evidence from Primary School," *Journal of Human Resources*, 2013, *48* (1), 236–264.

**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.

**Cunha, Flavio and James Heckman**, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

**Dee, Thomas S**, "Teachers, Race, and Student Achievement in a Randomized Experiment," *Review of Economics and Statistics*, 2004, *86* (1), 195–210.

_ , "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.

**Diprete, Thomas A and Claudia Buchmann**, "Gender-Specific Trends in the Value of Education and the Emerging Gender Gap in College Completion," *Demography*, 2006, *43* (1), 1–24.

**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.

**Fairlie, Robert W, Florian Hoffmann, and Philip Oreopoulos**, "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom," *The American Economic Review*, 2014, *104* (8), 2567–2591.

**Ferguson, Ronald F**, "Can Schools Narrow the Black-White Test Score Gap," *The Black-White Test Score Gap*, 1995, *318*, 374.

**Flossmann, Anton L, Rémi Piatek, and Laura Wichert**, "Going Beyond Returns to Education: The Role of Noncognitive Skills on Wages in Germany," Technical Report, Konstanz: Univ. 2008.

**Fortin, Nicole M, Philip Oreopoulos, and Shelley Phipps**, "Leaving Boys Behind Gender Disparities in High Academic Achievement," *Journal of Human Resources*, 2015, *50* (3), 549–579.

**Heckman, James J and Tim Kautz**, "Fostering and Measuring Skills: Interventions that Improve Character and Cognition," Technical Report, National Bureau of Economic Research 2013.

_ **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.

**Hoffmann, Florian and Philip Oreopoulos**, "A Professor Like Me: The Influence of Instructor Gender on College Achievement," *Journal of Human Resources*, 2009, *44* (2), 479–494.

**Holmlund, Helena and Krister Sund**, "Is the Gender Gap in School Performance Affected by the Sex of the Teacher?," *Labour Economics*, 2008, *15* (1), 37–53.

**Jacobs, Jerry A**, "Gender Inequality and Higher Education," *Annual Review of Sociology*, 1996, pp. 153–185.

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

**Jones, Susanne M and Kathryn Dindia**, "A Meta-Analytic Perspective on Sex Equity in the Classroom," *Review of Educational Research*, 2004, *74* (4), 443–471.

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

**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, pp. 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.

**Neumark, David and Rosella Gardecki**, "Women Helping Women? Role Model and Mentoring Effects on Female Ph. D. Students in Economics," *Journal of Human Resources*, 1998, *33* (1), 220–247.

**Nixon, Lucia A and Michael D Robinson**, "The Educational Attainment of Young Women: Role Model Effects of Female High School Faculty," *Demography*, 1999, *36* (2), 185–194.

**Rothstein, Donna S**, "Do Female Faculty Influence Female Students' Educational and Labor Market Attainments?," *Industrial & Labor Relations Review*, 1995, *48* (3), 515–530.

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

**Spencer, Steven J, Claude M Steele, and Diane M Quinn**, "Stereotype Threat and Women's Math Performance," *Journal of Experimental Social Psychology*, 1999, *35* (1), 4–28.

**Steele, Claude M**, "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance.," *American Psychologist*, 1997, *52* (6), 613.

_ **and Joshua Aronson**, "Stereotype Threat and the Intellectual Test Performance of African Americans," *Journal of Personality and Social Psychology*, 1995, *69* (5), 797.

**Vincent-Lancrin, Stéphan**, "The Reversal of Gender Inequalities in Higher Education: An On-Going Trend," 2008.

**Winters, Marcus A, Robert C Haight, Thomas T Swaim, and Katarzyna A Pickering**, "The Effect of Same-Gender Teacher Assignment on Student Achievement in the Elementary and Secondary Grades: Evidence from Panel Data," *Economics of Education Review*, 2013, *34*, 69–75.

Table 1. Summary Statistics

| *Panel A. Outcome Variables* | | | |
|---|---|---|---|
| | Mean | Standard Deviation | Observation |
| *Academic Outcomes:* | (1) | (2) | (3) |
| Test Score | 81.26 | 28.34 | 25783 |
| Self-Assessment Score | 2.48 | 0.91 | 26301 |
| *Noncognitive Outcomes:* | | | |
| Depressed | 2.24 | 1.00 | 8772 |
| Blue | 1.98 | 1.06 | 8743 |
| Unhappy | 2.28 | 1.05 | 8762 |
| Pessimistic | 1.75 | 1.07 | 8734 |
| School Life is Fulfilling | 3.38 | 0.86 | 8852 |
| Confident about Future | 3.26 | 0.72 | 8924 |
| Social Activities: Public Enrichment | 2.02 | 1.04 | 8686 |
| Social Activities: Private Recreation | 2.44 | 1.28 | 8653 |
| *Panel B. Regressors of Interest* | | | |
| Female Student | 0.49 | 0.50 | 8910 |
| Female Head Teacher | 0.65 | 0.48 | 8988 |
| Female Subject Teacher | 0.78 | 0.41 | 19550 |
| *Panel C. Student Pre-Determined Variables* | | | |
| Age | 13.94 | 1.35 | 8815 |
| Minority | 0.11 | 0.31 | 8968 |
| Local Residence | 0.80 | 0.40 | 8811 |
| Only Child | 0.51 | 0.50 | 8986 |
| Attend Kindergarten | 0.82 | 0.39 | 8912 |
| Repeat in Primary School | 0.11 | 0.32 | 8988 |
| Academic Ranking in Primary School | 15.65 | 11.82 | 8121 |
| Mother's Education | 10.07 | 3.63 | 8966 |
| Father's Education | 10.73 | 3.28 | 8966 |

## Table 2. Balancing Test

| | Female Head Teacher | Male Head Teacher | Difference | Conditional Difference |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Panel A. Student Characteristics* | | | | |
| Female | 0.50 | 0.47 | 0.029** | 0.020 |
| | [0.09] | [0.1] | (0.013) | (0.014) |
| Age | 13.87 | 14.16 | -0.292* | -0.195 |
| | [1.14] | [1.22] | (0.173) | (0.199) |
| Minority | 0.07 | 0.15 | -0.082*** | -0.007 |
| | [0.12] | [0.28] | (0.028) | (0.007) |
| Local Residence | 0.77 | 0.83 | -0.060** | -0.004 |
| | [0.2] | [0.21] | (0.030) | (0.022) |
| Only Child | 0.56 | 0.39 | 0.171*** | 0.027 |
| | [0.25] | [0.31] | (0.040) | (0.017) |
| Attend Kindergarten | 0.84 | 0.77 | 0.066*** | 0.012 |
| | [0.11] | [0.16] | (0.019) | (0.014) |
| Repeat in Primary School | 0.09 | 0.17 | -0.077*** | -0.007 |
| | [0.13] | [0.17] | (0.021) | (0.011) |
| Academic Ranking in Primary School | 14.67 | 15.96 | -1.293** | -0.394 |
| | [3.64] | [4.04] | (0.560) | (0.538) |
| Mother's Education | 10.49 | 9.23 | 1.261*** | 0.103 |
| | [2.04] | [2.58] | (0.331) | (0.130) |
| Father's Education | 10.96 | 10.08 | 0.882*** | 0.199 |
| | [2.02] | [2.03] | (0.301) | (0.123) |
| *Panel B. Classroom and Teacher Characteristics* | | | | |
| Proportion Other Female Subject Teachers | 0.81 | 0.77 | 0.033 | -0.078 |
| | [0.28] | [0.31] | (0.044) | (0.049) |
| Chinese Teacher's Teaching Experience | 16.52 | 16.43 | 0.094 | -0.829 |
| | [8.77] | [9.58] | (1.524) | (2.062) |
| Math Teacher's Teaching Experience | 17.05 | 16.66 | 0.389 | -0.965 |
| | [9.26] | [9.39] | (1.710) | (2.356) |
| English Teacher's Teaching Experience | 15.12 | 14.21 | 0.909 | -1.636 |
| | [9.34] | [9.3] | (1.565) | (1.952) |
| Class Size | 41.39 | 47.04 | -5.655*** | -0.115 |
| | [12.32] | [15.93] | (2.016) | (1.321) |

*Notes:* Column (1) and column (2) present the mean and standard deviation of the student predetermined variables. Column (3) and column (4) present the difference and standard error. Specifically, in column (3) and column (4), each cell represents a separate regression, in which the independent variable is an indicator for female head teacher, and the dependent variable is the corresponding student, classroom and teacher characteristic as listed above. All specifications in column (4) include school fixed effects. Standard errors are reported in parentheses. ***significant at the 1% level, **5% level, *10% level.

Table 3. Teacher Gender Effect on Students' Academic Outcomes

| | Test Score | | Self-Assessment Score | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Female Teacher × Female Student | 0.141** | 0.144*** | 0.269*** | 0.273*** |
| | (0.055) | (0.053) | (0.051) | (0.052) |
| Female Teacher | -0.035 | -0.040 | -0.115*** | -0.117*** |
| | (0.038) | (0.037) | (0.032) | (0.032) |
| Female Student | 0.316*** | 0.194*** | -0.011 | -0.062 |
| | (0.051) | (0.048) | (0.046) | (0.046) |
| Subject Fixed Effects | Yes | Yes | Yes | Yes |
| School Fixed Effects | Yes | Yes | Yes | Yes |
| Student Control | No | Yes | No | Yes |
| Observations | 16,355 | 16,355 | 16,733 | 16,733 |
| R-squared | 0.053 | 0.256 | 0.015 | 0.060 |

*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. Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and whether repeated a grade in primary school. ***significant at the 1% level, **5% level, *10% level.

Table 4A. Teacher Gender Effect on Noncognitive Outcomes - Mental Stress

| | Depressed | Blue | Unhappy | Pessimistic | AES | Depressed | Blue | Unhappy | Pessimistic | AES |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Female Teacher × | -0.155*** | -0.168*** | -0.135*** | -0.036 | -0.123*** | -0.158*** | -0.174*** | -0.138*** | -0.038 | -0.127*** |
| Female Student | (0.054) | (0.055) | (0.052) | (0.049) | (0.044) | (0.053) | (0.055) | (0.052) | (0.048) | (0.043) |
| Female Teacher | -0.029 | -0.035 | -0.043 | -0.072 | | -0.009 | -0.012 | -0.023 | -0.054 | |
| | (0.048) | (0.048) | (0.047) | (0.045) | | (0.045) | (0.043) | (0.043) | (0.042) | |
| Female Student | 0.253*** | 0.069 | 0.128*** | -0.046 | | 0.277*** | 0.092* | 0.157*** | -0.014 | |
| | (0.044) | (0.047) | (0.044) | (0.040) | | (0.043) | (0.047) | (0.043) | (0.039) | |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Student Control | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,288 | 7,288 | 7,288 | 7,288 | 7,288 | 7,288 | 7,288 | 7,288 | 7,288 | 7,288 |
| R-squared | 0.044 | 0.046 | 0.042 | 0.033 | | 0.044 | 0.046 | 0.042 | 0.033 | |

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

Table 4B. Teacher Gender Effect on Noncognitive Outcomes - Social Acclimation and Satisfaction

| | Fulfilling of Life | Confident abt Future | Social Activities with Peers | | AES | Fulfilling of Life | Confident abt Future | Social Activities with Peers | | AES |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Public Enrichment | Private Recreation | | | | Public Enrichment | Private Recreation | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Female Teacher × | 0.069 | 0.107* | 0.032 | 0.139*** | 0.087*** | 0.075 | 0.113** | 0.035 | 0.137*** | 0.091*** |
| Female Student | (0.052) | (0.057) | (0.044) | (0.047) | (0.031) | (0.050) | (0.055) | (0.043) | (0.047) | (0.030) |
| Female Teacher | 0.087 | 0.058 | 0.090* | -0.005 | | 0.066 | 0.027 | 0.064 | -0.016 | |
| | (0.055) | (0.050) | (0.054) | (0.044) | | (0.052) | (0.044) | (0.048) | (0.042) | |
| Female Student | 0.029 | -0.105** | 0.002 | -0.092** | | 0.003 | -0.163*** | -0.027 | -0.108*** | |
| | (0.043) | (0.050) | (0.036) | (0.040) | | (0.042) | (0.048) | (0.035) | (0.039) | |
| School Fixed | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Student Control | No | No | No | No | No | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,141 | 7,141 | 7,141 | 7,141 | 7,141 | 7,141 | 7,141 | 7,141 | 7,141 | 7,141 |
| R-squared | 0.070 | 0.067 | 0.196 | 0.211 | | 0.083 | 0.110 | 0.222 | 0.229 | |

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

Table 4C. Teacher Gender Effect on Noncognitive Outcomes, Controlling for Test Score

| | Mental Stress | | | | | Social Acclimation and Satisfaction | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Depressed | Blue | Unhappy | Pessimistic | AES | Fulfilling of Life | Confident abt Future | Social Activities with Peers: Public Enrichment | Social Activities with Peers: Private Recreation | AES |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Female Teacher × | -0.158*** | -0.173*** | -0.138*** | -0.036 | -0.126*** | 0.072 | 0.109** | 0.036 | 0.139*** | 0.089*** |
| Female Student | (0.053) | (0.055) | (0.052) | (0.048) | -0.043 | (0.050) | (0.055) | (0.043) | (0.046) | (0.030) |
| Female Teacher | -0.007 | -0.008 | -0.023 | -0.049 | | 0.025 | -0.016 | 0.078* | -0.001 | |
| | (0.045) | (0.044) | (0.043) | (0.042) | | (0.049) | (0.042) | (0.047) | (0.042) | |
| Female Student | 0.281*** | 0.105** | 0.158*** | 0.001 | | -0.028 | -0.195*** | -0.017 | -0.097** | |
| | (0.044) | (0.047) | (0.043) | (0.039) | | (0.042) | (0.048) | (0.035) | (0.039) | |
| Average Test Score | -0.016 | -0.047*** | -0.004 | -0.055*** | | 0.101*** | 0.122*** | -0.023 | -0.041*** | |
| | (0.016) | (0.015) | (0.016) | (0.016) | | (0.016) | (0.017) | (0.015) | (0.015) | |
| School Fixed | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Student Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,288 | 7,288 | 7,288 | 7,288 | 7,288 | 7,141 | 7,141 | 7,141 | 7,141 | 7,141 |
| R-squared | 0.044 | 0.046 | 0.042 | 0.033 | | 0.089 | 0.120 | 0.222 | 0.230 | 0.089 |

*Notes:* Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and whether repeated a grade in primary school. Average Test score is the normalized student's average test score in three core subjects. Standard errors are clustered at class level. ***significant at the 1% level, **5% level, *10% level.

Table 5. Mechanism: Is it Due to Teacher Gender or Other Teacher Characteristics?

| | Academic Outcome | | Mental Stress | | | | | Social Acclimation and Satisfaction | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Test score | Self-Assessment | Depressed | Blue | Unhappy | Pessimistic | AES | School Fulfilling | Confident abt Future | Social Activities with Peers | | AES |
| | | | | | | | | | | Public Enrichment | Private Recreation | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Female Teacher × | 0.121** | 0.276*** | -0.165*** | -0.209*** | -0.160*** | -0.021 | -0.140*** | 0.106* | 0.091 | 0.018 | 0.136** | 0.089*** |
| Female Student | (0.054) | (0.055) | (0.056) | (0.057) | (0.057) | (0.051) | (0.046) | (0.054) | (0.057) | (0.048) | (0.053) | (0.033) |
| 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 |
| Student Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Female Student × Teacher Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Female Teacher × Teacher Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 15,655 | 16,023 | 7,075 | 7,075 | 7,075 | 7,075 | 7,075 | 6,930 | 6,930 | 6,930 | 6,930 | 6,930 |
| R-squared | 0.258 | 0.062 | 0.063 | 0.063 | 0.063 | 0.050 | | 0.089 | 0.114 | 0.229 | 0.235 | |

*Notes:* Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and whether repeated a grade in primary school. Teacher control includes: teacher's age, education, teaching experience, working experience as head teacher (for head teacher only) and dummy variables indicating married, has a certificated credential, graduated from normal college, and teach a main subject (for head teacher only). Standard errors are clustered at class level. ***significant at the 1% level, **5% level, *10% level.

## Table 6. Mechanism: Teachers' Behaviors

*Panel A: Head Teachers' Behaviors*

| | Outcome: Head Teachers' Praises | | | Outcome: Head Teachers' Criticisms | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Female Teacher × | 0.062 | 0.066 | 0.062 | -0.130*** | -0.134*** | -0.127*** |
| Female Student | (0.051) | (0.051) | (0.051) | (0.050) | (0.049) | (0.048) |
| Female Teacher | 0.123** | 0.099* | 0.088 | 0.100** | 0.111** | 0.125*** |
| | (0.058) | (0.057) | (0.056) | (0.047) | (0.048) | (0.046) |
| Female Student | 0.044 | -0.011 | -0.042 | -0.236*** | -0.208*** | -0.167*** |
| | (0.042) | (0.043) | (0.042) | (0.042) | (0.041) | (0.040) |
| School Fixed | Yes | Yes | Yes | Yes | Yes | Yes |
| Student Control | No | Yes | Yes | No | Yes | Yes |
| Test Score Control | No | No | Yes | No | No | Yes |
| Observations | 7,440 | 7,440 | 7,440 | 7,444 | 7,444 | 7,444 |
| R-squared | 0.095 | 0.128 | 0.138 | 0.051 | 0.060 | 0.075 |

*Panel B: Subject Teachers' Behaviors*

| | Outcome: Subject Teachers' Class Questionings | | | Outcome: Subject Teachers' Praises | | |
|---|---|---|---|---|---|---|
| Female Teacher × | 0.129*** | 0.126*** | 0.111*** | 0.100** | 0.098** | 0.080** |
| Female Student | (0.042) | (0.042) | (0.041) | (0.039) | (0.039) | (0.038) |
| Female Teacher | -0.015 | -0.012 | -0.007 | 0.016 | 0.017 | 0.022 |
| | (0.048) | (0.048) | (0.047) | (0.042) | (0.043) | (0.043) |
| Female Student | -0.087** | -0.123*** | -0.144*** | -0.034 | -0.077** | -0.101*** |
| | (0.039) | (0.039) | (0.038) | (0.036) | (0.037) | (0.037) |
| Subject Fixed | Yes | Yes | Yes | Yes | Yes | Yes |
| School Fixed | Yes | Yes | Yes | Yes | Yes | Yes |
| Student Control | No | Yes | Yes | No | Yes | Yes |
| Test Score Control | No | No | Yes | No | No | Yes |
| Observations | 16,273 | 16,273 | 16,273 | 16,275 | 16,275 | 16,275 |
| R-squared | 0.085 | 0.103 | 0.111 | 0.100 | 0.122 | 0.134 |

*Notes:* Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and whether repeated a grade in primary school. In panel A, test score control is the student's average test score in three core subjects. In panel B, test score control is the student's test score in the corresponding subject. Standard errors are clustered at class level. ***significant at the 1% level, **5% level, *10% level.

## Table 7. Mechanism: Students' Responses

*Panel A. Reduced Gender Stereotype Threats*

| | Outcome: Boys are More Talented in Learning Math | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Female Math Teacher × Female Student | -0.141* | -0.140* | -0.139* |
| | (0.079) | (0.079) | (0.079) |
| Female Teacher | 0.074 | 0.069 | 0.068 |
| | (0.057) | (0.057) | (0.058) |
| Female Student | -0.113* | -0.117* | -0.118* |
| | (0.067) | (0.067) | (0.067) |
| School Fixed Effects | Yes | Yes | Yes |
| Student Control | No | Yes | Yes |
| Test Score Control | No | No | Yes |
| Observations | 5,233 | 5,233 | 5,233 |
| R-squared | 0.059 | 0.061 | 0.061 |

*Panel B. Role Model Effects*

| | Outcome: The Subject is Useful for My Future | | |
|---|---|---|---|
| Female Subject Teacher × Female Student | 0.202*** | 0.196*** | 0.173*** |
| | (0.043) | (0.042) | (0.041) |
| Female Teacher | -0.072* | -0.057 | -0.050 |
| | (0.043) | (0.040) | (0.040) |
| Female Student | -0.002 | -0.040 | -0.070* |
| | (0.040) | (0.039) | (0.039) |
| Subject Fixed Effects | Yes | Yes | Yes |
| School Fixed Effects | Yes | Yes | Yes |
| Student Control | No | Yes | Yes |
| Test Score Control | No | No | Yes |
| Observations | 16,296 | 16,296 | 16,296 |
| R-squared | 0.048 | 0.079 | 0.096 |

*Notes:* Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and whether repeated a grade in primary school. In panel A, test score control is the student's test score in math. In panel B, test score control is the student's test score in the corresponding subject. Standard errors are clustered at class level. ***significant at the 1% level, **5% level, *10% level.

Table 8. Heterogeneous Effects: Test Score

| | Mother's Education >9 years | Mother's Education <= 9 years | Ethnic Majority | Ethnic Minority | Living with Parents | Parents Migrated Out |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Female Teacher × | 0.035 | 0.217*** | 0.088* | 0.421*** | 0.098* | 0.296*** |
| Female Student | (0.068) | (0.065) | (0.051) | (0.155) | (0.054) | (0.096) |
| Female Teacher | 0.023 | -0.075 | -0.004 | -0.210** | -0.004 | -0.154** |
| | (0.047) | (0.048) | (0.037) | (0.094) | (0.039) | (0.061) |
| Female Student | 0.213*** | 0.185*** | 0.245*** | -0.001 | 0.240*** | 0.052 |
| | (0.063) | (0.059) | (0.047) | (0.114) | (0.050) | (0.091) |
| Subject Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Student Control | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6,938 | 9,417 | 14,795 | 1,560 | 13,343 | 3,012 |
| R-squared | 0.285 | 0.247 | 0.269 | 0.245 | 0.263 | 0.252 |

*Notes:* Test score is normalized by subject and school grade. In our estimation sample, the medium of mother's education is 9 years. Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten and whether repeated a grade in primary school. Standard errors are clustered at class level. ***significant at the 1% level, **5% level, *10% level.

## Table 9. Heterogeneous Effects: Noncognitive Outcomes

*Panel A. Mental Stress*

| | Depressed | Blue | Unhappy | Pessimistic | AES | Depressed | Blue | Unhappy | Pessimistic | AES |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mother's Education > 9 years | | | | | Mother's Education <=9 years | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Female Teacher × | -0.108 | -0.129 | -0.105 | 0.004 | -0.082 | -0.154** | -0.175** | -0.128** | -0.029 | -0.126** |
| Female Student | (0.083) | (0.082) | (0.086) | (0.065) | (0.066) | (0.069) | (0.071) | (0.064) | (0.063) | (0.053) |
| Observations | 3,136 | 3,132 | 3,133 | 3,133 | 3,122 | 4,426 | 4,409 | 4,419 | 4,397 | 4,372 |
| R-squared | 0.062 | 0.068 | 0.061 | 0.072 | | 0.069 | 0.065 | 0.060 | 0.049 | |
| | Ethnic Majority | | | | | Ethnic Minority | | | | |
| Female Teacher × | -0.122** | -0.147*** | -0.101* | 0.018 | -0.088** | -0.161 | -0.228 | -0.284* | -0.402** | -0.279** |
| Female Student | (0.052) | (0.056) | (0.053) | (0.052) | (0.044) | (0.173) | (0.214) | (0.170) | (0.155) | (0.135) |
| Observations | 6,775 | 6,756 | 6,768 | 6,753 | 6,722 | 787 | 785 | 784 | 777 | 772 |
| R-squared | 0.054 | 0.055 | 0.053 | 0.049 | | 0.143 | 0.143 | 0.119 | 0.088 | |
| | Living with Parents | | | | | Parents Migrated Out | | | | |
| Female Teacher × | -0.127** | -0.138** | -0.096* | -0.007 | -0.093* | -0.213 | -0.255* | -0.248** | -0.153 | -0.223** |
| Female Student | (0.058) | (0.060) | (0.057) | (0.055) | (0.048) | (0.130) | (0.132) | (0.115) | (0.109) | (0.094) |
| Observations | 6,130 | 6,111 | 6,120 | 6,107 | 6,076 | 1,432 | 1,430 | 1,432 | 1,423 | 1,418 |
| R-squared | 0.057 | 0.060 | 0.052 | 0.047 | | 0.097 | 0.106 | 0.101 | 0.084 | |

## Table 9. Heterogeneous Effects: Noncognitive Outcomes (Continued)

*Panel B. Social Acclimation and Satisfaction*

| | School is Fulfilling | Confident abt Future | Social Activities with Peers | | AES | School is Fulfilling | Confident abt Future | Social Activities with Peers | | AES |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Public Enrichment | Private Recreation | | | | Public Enrichment | Private Recreation | |
| | Mother's Education>9 years | | | | | Mother's Education<=9 years | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Female Teacher × | 0.149** | 0.066 | -0.024 | 0.166* | 0.099* | 0.064 | 0.098 | 0.055 | 0.119** | 0.087** |
| Female Student | (0.071) | (0.069) | (0.077) | (0.089) | (0.052) | (0.064) | (0.065) | (0.053) | (0.0481) | (0.035) |
| Observations | 3,171 | 3,177 | 3,105 | 3,096 | 3,058 | 4,458 | 4,504 | 4,395 | 4,375 | 4,289 |
| R-squared | 0.066 | 0.101 | 0.116 | 0.128 | | 0.096 | 0.100 | 0.218 | 0.186 | |
| | Ethnic Majority | | | | | Ethnic Minority | | | | |
| Female Teacher × | 0.139*** | 0.086 | 0.031 | 0.142*** | 0.102*** | 0.035 | 0.107 | -0.045 | -0.078 | 0.016 |
| Female Student | (0.051) | (0.053) | (0.046) | (0.049) | (0.033) | (0.137) | (0.158) | (0.130) | (0.124) | (0.072) |
| Observations | 6,848 | 6,886 | 6,727 | 6,700 | 6,597 | 781 | 795 | 773 | 771 | 750 |
| R-squared | 0.075 | 0.110 | 0.203 | 0.215 | | 0.180 | 0.148 | 0.359 | 0.328 | |
| | Living with Parents | | | | | Parents Migrated Out | | | | |
| Female Teacher × | 0.088 | 0.104* | 0.001 | 0.164*** | 0.090** | 0.186 | 0.034 | 0.115 | 0.032 | 0.108* |
| Female Student | (0.054) | (0.060) | (0.048) | (0.051) | (0.036) | (0.116) | (0.110) | (0.081) | (0.094) | (0.059) |
| Observations | 6,182 | 6,222 | 6,087 | 6,070 | 5,970 | 1,447 | 1,459 | 1,413 | 1,401 | 1,377 |
| R-squared | 0.081 | 0.110 | 0.213 | 0.222 | | 0.140 | 0.140 | 0.250 | 0.251 | |
| School fixed | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Student control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

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

## Appendix Tables

### Table A1. Gender Difference across Classes with Female and Male Head Teachers

| | Age | Minority | Local Residence | Only Child | Attend Kindergarten | Repeat in Primary School | Academic Ranking in Primary School | Mother's Education | Father's Education |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Female Teacher × | 0.026 | 0.020** | 0.003 | -0.033 | -0.009 | -0.009 | 0.264 | 0.176 | 0.087 |
| Female Student | (0.048) | (0.009) | (0.014) | (0.021) | (0.017) | (0.012) | (0.605) | (0.110) | (0.117) |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 8,742 | 8,891 | 8,735 | 8,908 | 8,835 | 8,910 | 8,054 | 8,888 | 8,888 |
| R-squared | 0.398 | 0.513 | 0.222 | 0.303 | 0.094 | 0.187 | 0.091 | 0.406 | 0.364 |

*Notes:* Each cell represents a separate regression in which the dependent variable is the student pre-determined characteristic as listed above, and the independent variable is interaction term between head teacher's gender and student's gender, dummy for head teacher's gender, and dummy for student's gender. School fixed effects are included. Standard errors are clustered at class level. ***significant at the 1% level, **5% level, *10% level.

Table A2. Robustness Check: Female Head Teacher and Academic Outcomes

| | Test Score | | Self-Assessment Score | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Female Subject Teacher × | 0.150*** | 0.151*** | 0.277*** | 0.280*** |
| Female Student | (0.056) | (0.053) | (0.052) | (0.052) |
| Female Subject Teacher | -0.040 | -0.044 | -0.119*** | -0.121*** |
| | (0.038) | (0.036) | (0.033) | (0.032) |
| Female Student | 0.342*** | 0.214*** | 0.014 | -0.038 |
| | (0.055) | (0.054) | (0.047) | (0.048) |
| Female Head Teacher × | -0.053 | -0.040 | -0.050 | -0.046 |
| Female Student | (0.047) | (0.044) | (0.039) | (0.040) |
| Female Head Teacher | 0.139*** | 0.101** | 0.086*** | 0.065** |
| | (0.050) | (0.045) | (0.033) | (0.032) |
| Subject Fixed Effects | Yes | Yes | Yes | Yes |
| School Fixed Effects | Yes | Yes | Yes | Yes |
| Student Control | No | Yes | No | Yes |
| Observations | 16,355 | 16,355 | 16,733 | 16,733 |
| R-squared | 0.054 | 0.257 | 0.015 | 0.060 |

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

## Table A3. Robustness Check: Alternative Measurement of Teacher Gender

| | Mental Stress | | | | | Social Acclimation and Satisfaction | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Depressed | Blue | Unhappy | Pessimistic | AES | School is Fulfilling | Confident abt Future | Social Activities | | AES |
| | | | | | | | | Public Enrichment | Private Recreation | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Proportion Female Teacher | -0.358*** | -0.322*** | -0.291*** | -0.013 | -0.246*** | 0.080 | 0.297*** | -0.005 | 0.274*** | 0.162*** |
| × Female Student | (0.116) | (0.114) | (0.102) | (0.090) | (0.086) | (0.094) | (0.096) | (0.096) | (0.079) | (0.047) |
| Proportion Female Teacher | 0.059 | -0.093 | -0.010 | -0.125 | | 0.311*** | 0.152 | 0.181 | -0.139 | |
| | (0.130) | (0.131) | (0.139) | (0.120) | | (0.107) | (0.095) | (0.114) | (0.087) | |
| Female Student | 0.437*** | 0.232*** | 0.266*** | -0.006 | | 0.009 | -0.296*** | 0.023 | -0.198*** | |
| | (0.089) | (0.088) | (0.078) | (0.069) | | (0.071) | (0.077) | (0.073) | (0.062) | |
| School Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Student Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 4,342 | 4,342 | 4,342 | 4,342 | 4,342 | 4,267 | 4,267 | 4,267 | 4,267 | 4,267 |
| R-squared | 0.070 | 0.066 | 0.073 | 0.056 | | 0.106 | 0.114 | 0.221 | 0.241 | |

*Notes:* "Proportion female teacher" is an alternative measurement of teacher gender, and it is the time weighted proportion of female teachers in class. Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attended kindergarten, and whether repeated a grade in primary school. Standard errors are clustered at class level. ***significant at the 1% level, **5% level, *10% level.

Table A4. Robustness Check: Including Teacher Control, Grade Fixed Effect, and Student Fixed Effect

| | Test Score | | | Self-Assessment Score | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Female Teacher × Female Student | 0.168*** | 0.167*** | 0.230** | 0.287*** | 0.292*** | 0.295*** |
| | (0.053) | (0.053) | (0.093) | (0.051) | (0.053) | (0.112) |
| Female Teacher | -0.053 | -0.044 | -0.086 | -0.167*** | -0.125*** | -0.114 |
| | (0.037) | (0.038) | (0.055) | (0.046) | (0.034) | (0.073) |
| Female Student | 0.170*** | 0.171*** | NA | -0.080* | -0.077* | NA |
| | (0.048) | (0.048) | | (0.045) | (0.046) | |
| 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 |
| Student Fixed Effects | No | No | Yes | No | No | Yes |
| Student Control | Yes | Yes | NA | Yes | Yes | NA |
| Teacher Control | No | Yes | Yes | No | Yes | Yes |
| Observations | 15,655 | 15,655 | 15,655 | 16,023 | 16,023 | 16,023 |
| R-squared | 0.256 | 0.257 | 0.829 | 0.130 | 0.061 | 0.655 |

*Notes:* Test score and self-assessment score are normalized by subject, grade and school. Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attend kindergarten, and whether repeated a grade in primary school. 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. ***significant at the 1% level, **5% level, *10% level.

Table A5. Robustness Check: Including Teacher Control and Grade Fixed Effect

| | Mental Stress | | | | | Social Acclimation and Satisfaction | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Social Activities with peers | | |
| | Depressed | Blue | Unhappy | Pessimistic | AES | School is Fulfilling | Confident about Future | Public Enrichment | Private Recreation | AES |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Female Teacher × | -0.168*** | -0.186*** | -0.134** | -0.030 | -0.130*** | 0.098* | 0.127** | 0.012 | 0.152*** | 0.097*** |
| Female Student | (0.057) | (0.054) | (0.054) | (0.048) | (0.044) | (0.052) | (0.055) | (0.044) | (0.046) | (0.030) |
| Female Teacher | -0.017 | -0.019 | -0.033 | -0.045 | | 0.045 | 0.047 | 0.085 | -0.023 | |
| | (0.046) | (0.045) | (0.046) | (0.045) | | (0.057) | (0.044) | (0.054) | (0.042) | |
| Female Student | 0.279*** | 0.100** | 0.146*** | -0.026 | | -0.006 | -0.159*** | -0.011 | -0.121*** | |
| | (0.047) | (0.046) | (0.045) | (0.039) | | (0.043) | (0.048) | (0.036) | (0.039) | |
| 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 |
| Student Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,075 | 7,075 | 7,075 | 7,075 | 7,075 | 6,930 | 6,930 | 6,930 | 6,930 | 6,930 |
| R-squared | 0.059 | 0.061 | 0.061 | 0.049 | | 0.088 | 0.113 | 0.228 | 0.234 | |

*Notes:* Test score and self-assessment score are normalized by subject, grade and school. Student control includes: student's age, academic ranking in primary school, mother's education, father's education, and dummy variables indicating minority, local residence, only child in family, whether attend kindergarten, and whether repeated a grade in primary school. 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. ***significant at the 1% level, **5% level, *10% level.

## Table A6. Robustness Check: Sample Attrition

| | Academic Outcomes | | Mental Stress | | | | Social Acclimation and Satisfaction | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | Social Activities with Peers | |
| | Test Score | Self-Assessment Score | Depressed | Blue | Unhappy | Pessimistic | School is Fulfilling | Confident abt Future | Public Enrichment | Private Recreation |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Female Teacher | 0.002 | -0.006 | 0.021 | 0.025 | 0.023 | 0.025 | 0.026 | 0.018 | 0.018 | 0.015 |
| | (0.011) | (0.010) | (0.019) | (0.019) | (0.019) | (0.019) | (0.018) | (0.019) | (0.018) | (0.018) |
| School Fixed | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 19,389 | 19,389 | 8,988 | 8,988 | 8,988 | 8,988 | 8,988 | 8,988 | 8,988 | 8,988 |
| R-squared | 0.082 | 0.079 | 0.078 | 0.080 | 0.078 | 0.079 | 0.082 | 0.083 | 0.079 | 0.079 |

*Notes:* Female teacher indicates female subject teacher in column (1) and column (2), and indicates female head teacher from column (3) to column (10). Each cell represents a separate regression in which the dependent variable is the sample attrition dummy for each outcome as listed above, and the independent variable is a dummy for female teacher. Standard errors are clustered at class level. ***significant at the 1% level, **5% level, *10% level.