# The effect of female teachers on female students' academic performance and study behaviors: Evidence from Chinese middle schools*

**Lanxuan Xing** [i]

*Hanyang University*

**Eleanor Jawon Choi** [ii]

*Hanyang University*

## Abstract

This study examines the impact of female teachers on female students' academic performance and study behaviors exploiting random student-teacher assignment in Chinese middle schools. Analyzing data from the China Education Panel Survey, we find that female students achieve substantially higher test scores in mathematics and Chinese and exhibit improved class attendance behaviors when taught by a female teacher. The effect on math scores is more pronounced among disadvantaged or low-achieving students, whereas similar patterns are not evident in test scores for other subjects or in study behaviors.

*Key words: student-teacher gender match, test scores, study behaviors*

JEL Classification: I21, I24

---

* We thank Songman Kang, Jungmo Yoon, and brown bag seminar participants at Hanyang University for helpful comments and discussions. All errors are our own.

i) First author. College of Economics and Finance, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, South Korea. Email: xinglanxuan@hanyang.ac.kr

ii) Corresponding author. College of Economics and Finance, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, South Korea. Email: choiej@hanyang.ac.kr

First received, October 8, 2023; Revision received, November 12, 2023; Accepted, November 13, 2023

# 1 Introduction

There is a substantial and expanding body of literature examining the impact of student-teacher gender match on students' academic outcomes. Previous studies have found mixed results, with the direction and magnitude of the effect depending on the institutional and cultural context (Card et al., 2022). For instance, research tends to report greater benefits of having a female teacher for female students in settings where female role models are scarce.[1] Expanding upon prior research, this paper examines the effects of female teachers on academic performance and study behaviors of female students in China. Our study is within a context where strong gender stereotypes and persistent gender disparities prevail due to deeply ingrained patriarchal norms.

We analyze a nationally representative sample of seventh and ninth-grade students from the China Education Panel Survey (CEPS). To identify causal effects, we leverage on the unique institutional setting in Chinese middle schools, where students are randomly assigned to classrooms and teachers at the beginning of their seventh grade, and this assignment remains unchanged until the students graduate after completing the ninth grade. This research design allows us to avoid potential biases stemming from non-random student-teacher matching based on preferences.

By utilizing data from the first wave of CEPS conducted during the 2013–2014 academic year, our study focuses on the short-term effects of students being instructed by a female subject teacher for a period of less than three years after the student-teacher matches are formed. We assess students' academic performance through test scores in three core subjects: mathematics, Chinese, and English. To further investigate whether teacher gender can lead to behavioral changes related to students' engagement in learning, we examine class attendance, punctuality, average homework time, and participation in private out-of-school education for each core subject.

[1] Reviewing 24 articles published in economics journals since 2000, Card et al. (2022) observe that the positive impact of female teachers on the academic achievement of female students tends to be more pronounced in countries with greater gender disparities, including China, India, South Korea, and various African nations. In countries such as the US, Chile, Netherlands, Sweden, or Germany, where there is minimal or no discernible effect of female teachers on test scores of female students in general, the benefits are more distinct in STEM education, where women are in the minority.

As random classroom assignment is implemented for each entering cohort of students within schools, we estimate a linear regression model that controls for school-by-grade fixed effects. The treatment variables in the regression analysis consist of a female student dummy, a female teacher dummy, and the interaction between these two indicator variables. In this specification, the key parameter of interest is the coefficient on the interaction term, which represents the impact of a female teacher on female students relative to male students (or, equivalently, in comparison to male teachers). This empirical strategy has been widely employed in recent studies on teacher gender effects that exploit random student-teacher assignments (e.g., Carrell et al., 2010; Antecol et al., 2015; Lim and Meer, 2017, 2020; Gong et al., 2018; Xu and Li, 2018; Eble and Hu, 2020).

We verify the validity of the random classroom assignment design in two ways. First, we conduct balance tests and demonstrate that teacher gender in each subject is uncorrelated with students' background characteristics, conditional on school-by-grade fixed effects. Second, we also show the stability of coefficient estimates on teacher gender and student gender across specifications, with and without student and teacher characteristics, when school-by-grade fixed effects are included in the regression.

Our analysis shows that a female teacher enhances the academic performance of female students in the teacher's subject, with substantial test score gains, particularly in math and Chinese, amounting to approximately 0.2 standard deviations. Additionally, female students demonstrate a significant improvement in their class attendance behaviors, with a reduction in class absence and late attendance of 0.1–0.2 standard deviations when taught by a female teacher. From heterogeneity analyses and quantile regressions, we also find that the positive effect on math scores of having a female math teacher is substantially larger for disadvantaged or low-achieving students. However, we do not observe similar patterns in test scores for other subjects or in study behaviors.

This study is closely related to Gong et al. (2018), Xu and Li (2018), and Eble and Hu (2020), all of which investigate the impact of teacher gender on student outcomes using the CEPS data within the Chinese context. While these studies explore various aspects, such as noncognitive outcomes, self-perception, subjective or societal beliefs, in addition to test scores, our research specifically emphasizes behaviors related to class attendance and learning engagement. Furthermore, we conduct an in-depth analysis of how these effects vary by family background and across the distribution of test scores.

The remainder of the paper is organized as follows. Section 2 provides institutional background on classroom assignment procedures in Chinese middle schools and outlines the CEPS data. Section 3 describes the empirical strategy based on random classroom assignment and examines the internal validity of the research design. Section 4 presents the estimation results on students' test scores and study behaviors. Section 5 concludes.

## 2 Background and Data

In most Chinese middle schools, classroom and teacher assignments are made when students begin their seventh grade. This assignment remains unchanged until students graduate upon completing the ninth grade, with their classmate and teacher assignments staying constant over the three years. Students attend classes in their designated classrooms, where they receive instruction from subject teachers. Unlike in the US, the subject teachers, not students, rotate to the classrooms according to the class schedule. Throughout middle school, students primarily focus on three core subjects: mathematics, Chinese, and English. Additionally, they study subjects such as politics, history, biology, and geography. The educational system is structured in this way to provide students with a stable learning environment, facilitate the development of strong social connections, and ensure that students receive a comprehensive education.

To investigate the impact of gender matching between teachers and students on academic performance and study behaviors, we analyze data from the China Education Panel Survey (CEPS) conducted by the China Center for Survey and Data at Renmin University.[2] CEPS collects data on a nationally representative sample of seventh and ninth graders in middle schools, employing a stratified multi-stage sampling method in 28 randomly selected counties across China. Four schools are surveyed from each selected county. Within each school, two seventh-grade classes and two ninth-grade classes are randomly chosen, with all students in these selected classes included in the survey. The survey is administered to students, homeroom

[2] See http://ceps.ruc.edu.cn/English/Home.htm for more details about CEPS.

teachers, teachers specializing in three core subjects (math, Chinese, and English), school principals, and the parents of the students.

The first wave of CEPS was conducted during the 2013–2014 academic year.[3] Subsequent waves continue to track the students annually during middle school and intermittently after they graduate from middle school until 2043. However, data from only the first and second waves are currently accessible to the public. In this study, we focus on the short-term effects of student-teacher gender match using data from the first wave of CEPS. The survey's design enables us to examine academic outcomes within one year following classroom assignment for seventh graders and slightly over two years after classroom assignment for ninth graders.

We use the sample restriction criteria outlined in Gong et al. (2018) to ensure random student-teacher assignment within schools. Specifically, our analysis is restricted to students from schools that meet the following conditions: 1) the school principal confirms random classroom assignments for students, 2) class assignments remain unchanged for grades 8 and 9 after the initial allocation in seventh grade, and 3) teachers indicate that student assignment in the respective grade is not based on test scores. The second and third conditions are necessary for maintaining the random matching of ninth graders in the analysis sample with their teachers. Ultimately, our analysis sample comprises 9,901 students across 60 schools.

Table 1 presents summary statistics of key variables employed in the analysis. Panel A displays outcome variables, which include students' academic performance and study behaviors. Academic performance is measured by midterm exam scores in the three core subjects during the fall semester of 2013. We standardize the raw scores to have a mean of zero and a variance of one within school-by-grade groups. To evaluate students' study behavior, we use four variables related to class attendance, punctuality, average homework time, and participation in private out-of-school education for each core subject. Students are asked about their frequency of arriving late to class and skipping classes. The raw data is collected using a 4-point Likert scale (1 = completely disagree, 2 = somewhat disagree, 3 = somewhat agree, 4 = completely agree), and we standardize these raw variables within school-by-grade groups. Additionally, the mean hours students allocate to homework daily across a week are also standardized

[3] The school year in China begins in September.

within school-by-grade groups. We also utilize information on whether students engage in out-of-school private education, such as private tutoring, on math, Chinese, or English.

Table 1. Summary Statistics

| | Mean | Std. Dev. | N |
|---|---|---|---|
| ***Panel A. Outcomes*** | | | |
| ***Test scores*** | | | |
| Math | 0.000 | 0.994 | 9,629 |
| Chinese | 0.000 | 0.994 | 9,631 |
| English | -0.000 | 0.994 | 9,632 |
| ***Study behaviors*** | | | |
| Skip class (standardized) | -0.000 | 0.994 | 9,045 |
| Late to class (standardized) | -0.000 | 0.994 | 9,803 |
| Average homework time (standardized) | -0.000 | 0.994 | 9,253 |
| Private out-of-school education on math | 0.265 | 0.442 | 9,827 |
| Private out-of-school education on Chinese | 0.137 | 0.343 | 9,827 |
| Private out-of-school education on English | 0.264 | 0.441 | 9,827 |
| ***Panel B. Student characteristics*** | | | |
| Female student | 0.481 | 0.500 | 9,901 |
| Grade level | | | |
| 7th grade | 0.472 | 0.499 | 5,361 |
| 9th grade | 0.491 | 0.500 | 4,540 |
| Mothers' education | | | |
| Primary School or less | 0.252 | 0.434 | 9,869 |
| Middle School | 0.450 | 0.497 | 9,869 |
| High School | 0.152 | 0.359 | 9,869 |
| Bachelors' or more | 0.147 | 0.354 | 9,869 |
| Fathers' education | | | |
| Primary School or less | 0.157 | 0.364 | 9,869 |
| Middle School | 0.476 | 0.499 | 9,869 |
| High School | 0.193 | 0.395 | 9,869 |
| Bachelors' or more | 0.173 | 0.378 | 9,869 |
| Family financial status | | | |
| Poor | 0.192 | 0.394 | 9,859 |
| Moderate | 0.746 | 0.436 | 9,859 |
| Rich | 0.063 | 0.242 | 9,859 |
| Co-residence with parents | 0.780 | 0.414 | 9,901 |
| Only child | 0.455 | 0.498 | 9,901 |
| Father often gets drunk | 0.081 | 0.273 | 9,545 |
| Father and mother often quarrel | 0.098 | 0.297 | 9,544 |
| Parents have a good relationship | 0.835 | 0.371 | 9,623 |

| | | | |
|---|---|---|---|
| ***Panel C. Teacher characteristics*** | | | |
| ***Math*** | | | |
| Female teacher | 0.569 | 0.495 | 9,819 |
| Graduated from 4-year college | 0.507 | 0.500 | 9,627 |
| Have over 20 years of teaching experience | 0.343 | 0.475 | 9,901 |
| Major in education in college | 0.938 | 0.240 | 9,636 |
| Have administrative positions[a] | 0.406 | 0.491 | 9,221 |
| ***Chinese*** | | | |
| Female teacher | 0.762 | 0.426 | 9,668 |
| Graduated from 4-year college | 0.493 | 0.500 | 9,734 |
| Have over 20 years of teaching experience | 0.380 | 0.485 | 9,901 |
| Major in education in college | 0.940 | 0.238 | 9,734 |
| Have administrative positions[a] | 0.366 | 0.482 | 9,901 |
| ***English*** | | | |
| Female teacher | 0.908 | 0.289 | 9,716 |
| Graduated from 4-year college | 0.426 | 0.494 | 9,712 |
| Have over 20 years of teaching experience | 0.337 | 0.473 | 9,901 |
| Major in education in college | 0.913 | 0.282 | 9,646 |
| Have administrative positions[a] | 0.267 | 0.442 | 9,901 |

*Notes*. Test scores, skip class, late to class, average homework time are standardized to have zero mean and unit variance within school-by-grade groups. All student and teacher characteristics are binary variables.

[a] Administrative positions include head of the grade, head of the teaching and research group of a subject, teaching director or deputy teaching director, and deputy principal.

Panels B and C of Table 1 show the characteristics of students and teachers, respectively. Student and teacher gender serve as the key variables of interest in the analysis. Within the analyzed sample, 48% of students are female. Female teachers represent a substantial proportion across all three subjects: 57% of students have female math teachers, 76% have female Chinese teachers, and 91% have female English teachers. Other variables related to students' family backgrounds and teachers' characteristics are utilized as control variables in our empirical analysis. Appendix Table 1 provides detailed descriptions of each variable.

# 3 Empirical Strategy

To identify the effect of student-teacher gender matching on test scores and study behaviors, we leverage on the random assignment of students across classrooms and, as a result, across teachers within each grade and school. This type of research design has been employed in several related studies, including Carrell et al. (2010), Antecol et al. (2015), Lim and Meer (2017, 2020), Gong et al. (2018), and Xu and Li (2018). We estimate the typical regression model based on random student-teacher matches for each core subject:

$$Y_{ijgs} = \beta_1 S_i + \beta_2 T_j + \beta_3 S_i \times T_j + \boldsymbol{X}_{\mathrm{i}}'\boldsymbol{\gamma} + \boldsymbol{W}_j'\boldsymbol{\delta} + \theta_{gs} + \varepsilon_{ijgs}, \qquad (1)$$

where $Y_{ijgs}$ is the outcome for student $i$ assigned to teacher $j$ in grade $g$ at school $s$. As presented in panel A of Table 1, the outcome variables encompass exam scores and learning behaviors. $S_i$ and $T_j$ are binary variables indicating whether student $i$ is female and whether teacher $j$ is female, respectively. Thus, the interaction term, $S_i \times T_j$, indicates whether a female student is assigned to a female teacher for a specific subject. $\boldsymbol{X}_i$ and $\boldsymbol{W}_j$ represent vectors of student and teacher characteristics, excluding gender, which are listed in panels B and C of Table 1, respectively.[4] $\theta_{gs}$ denotes school-by-grade fixed effects. $\varepsilon_{ijgs}$ is an error term representing unobserved factors influencing the outcome. Standard errors are clustered at the school-by-grade level to account for correlations among students within the same grade in each school.

$\beta_1$ measures the mean difference in the outcome between male and female students when the teacher is male. $\beta_2$ represents the effect of a female teacher on the outcome for male students. $\beta_3$ captures the impact of a female teacher on the gender difference in the outcome, revealing whether female students experience greater benefits when paired with a teacher of the same gender. Note that $\beta_3$ is the primary parameter of interest in our analysis.

We examine the internal validity of the research design by testing the key identifying assumption of random classroom assignment within grades and schools in two ways. First, we conduct a balancing regression of teacher gender on students' background characteristics, conditional on school-by-grade fixed effects. Second, we check coefficient sensitivity across specifications with and without control variables on student and teacher characteristics

[4] Student controls include mothers' and fathers' education, family financial status, co-residence with parents, being an only child, father's drinking frequency, parental conflicts, and the quality of parental relationships. Teacher controls include the teacher's possession of a bachelor's degree from a four-year college, teaching experience exceeding 20 years, having majored in education at college, and holding administrative positions.

other than gender.

Table 2. Balance Tests

| | Female teacher | | |
|---|---|---|---|
| | **Math** | **Chinese** | **English** |
| | (1) | (2) | (3) |
| Female student | 0.010 | -0.002 | -0.000 |
| | (0.008) | (0.009) | (0.007) |
| Mothers' Education | | | |
| Middle school | 0.005 | -0.022 | -0.018 |
| | (0.018) | (0.017) | (0.012) |
| High school | -0.009 | -0.032 | -0.014 |
| | (0.020) | (0.020) | (0.012) |
| Bachelors' or more | 0.013 | -0.016 | -0.014 |
| | (0.032) | (0.016) | (0.012) |
| Fathers' Education | | | |
| Middle school | 0.008 | -0.019* | 0.006 |
| | (0.009) | (0.010) | (0.009) |
| High school | 0.013 | -0.019 | -0.002 |
| | (0.020) | (0.014) | (0.010) |
| Bachelors' or more | 0.018 | -0.029 | -0.007 |
| | (0.026) | (0.020) | (0.012) |
| Financial condition | | | |
| Moderate | 0.019 | -0.012 | 0.004 |
| | (0.015) | (0.017) | (0.011) |
| Rich | 0.042** | -0.004 | 0.028** |
| | (0.019) | (0.019) | (0.014) |
| Live together with parents | 0.012 | -0.020* | -0.006 |
| | (0.012) | (0.011) | (0.006) |
| Only child | 0.018* | -0.005 | -0.001 |
| | (0.010) | (0.010) | (0.007) |
| Father often gets drunk | -0.010 | -0.015 | 0.001 |
| | (0.017) | (0.016) | (0.013) |
| Father and mother often quarrel | 0.004 | -0.003 | -0.001 |
| | (0.008) | (0.007) | (0.009) |
| Parents have a good relationship | -0.002 | -0.005 | 0.007 |
| | (0.010) | (0.011) | (0.010) |
| $F$-statistic | 1.009 | 1.069 | 0.966 |
| | [0.449] | [0.392] | [0.492] |
| Adjusted $R^2$ | 0.665 | 0.739 | 0.509 |
| Observations | 9,310 | 9,169 | 9,221 |

*Notes*. Robust standard errors in parentheses are clustered at the school-by-grade level. $p$-values in brackets. *** $p$<0.01, ** $p$<0.05, * $p$<0.1

Table 2 performs a balance check, following the approach in Antecol et al. (2015) and Gong et al. (2018). In this exercise, we regress the gender of each subject teacher on student characteristics. If students are randomly assigned to teachers within each grade at each school, there should be no systematic association between teacher gender and observed student characteristics, conditional on school-by-grade fixed effects. The coefficient on each student characteristic is precisely zero, except for a few variables. Furthermore, when we perform a joint significance test, the null hypothesis of zero coefficients on all student characteristics cannot be rejected at any conventional level of significance (with $F$-statistics around one and $p$-values above 0.39) for all three subjects.

In Section 4, we also demonstrate that the coefficient estimates on student and teacher gender remain stable regardless of the addition of control variables, as long as school-by-grade fixed effects are included in the main regression (equation 1). Examining the movement of coefficients on the variables of interest after adding control variables in the regression is a common practice to assess the internal validity of selection on observables.[5] Given the findings from the two sets of validity check exercises, it is unlikely that our results are driven by non-random selection or endogenous matching of students and teachers.

# 4 Results

## 4.1 Effects on Test Scores

In Table 3, we first examine the effect of teacher gender on student academic performance in three core subjects: mathematics, Chinese and English. We provide coefficient estimates from the specification without control variables in odd-numbered columns and estimates from the specification with control variables in even-numbered columns. As both regression specifications yield very similar coefficient estimates on student and teacher gender variables, confirming the random classroom assignment, our subsequent focus is on the results obtained from the regression incorporating student and teacher controls.

5 See Altonji et al. (2005), Pei et al. (2019), and Oster (2019) for a detailed discussion of this method.

Table 3. Effects of Teacher Gender on Test Scores

| **Dependent variable:** | **Math** | | **Chinese** | | **English** | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Female student * Female teacher | 0.233*** | 0.227*** | 0.186** | 0.193** | 0.071 | 0.083 |
| | (0.072) | (0.068) | (0.076) | (0.074) | (0.092) | (0.084) |
| Female student | 0.012 | 0.005 | 0.479*** | 0.465*** | 0.527*** | 0.508*** |
| | (0.057) | (0.055) | (0.065) | (0.062) | (0.090) | (0.082) |
| Female teacher | 0.141 | 0.128 | -0.269 | -0.123 | -0.051 | 0.075 |
| | (0.162) | (0.197) | (0.220) | (0.216) | (0.121) | (0.114) |
| Controls | NO | YES | NO | YES | NO | YES |
| Adjusted $R^2$ | 0.006 | 0.040 | 0.085 | 0.121 | 0.075 | 0.125 |
| Observations | 9,548 | 9,548 | 9,404 | 9,404 | 9,467 | 9,467 |

*Notes*. Test scores are standardized to have zero mean and unit variance within school-by-grade groups. Control variables include student and teacher characteristics. See footnote 3 for a comprehensive list of control variables. We impute missing values of binary control variables with 0, and dummies for missing observations are also controlled for. Robust standard errors in parentheses are clustered at the school-by-grade level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

The first row of Table 3 presents the estimated effect on test scores when female teachers are matched with female students, showing our primary parameter of interest. The results indicate that a female subject teacher enhances the test scores of female students in the subject, relative to male students (or equivalently, when taught by male teachers). This benefit is substantial and statistically significant for math and Chinese, where the test score gains from a student-teacher gender match amount to about 0.2 standard deviations. For English, the point estimate is positive and non-negligible, although the effect size is less than half of that for the other two subjects and statistically indistinguishable from zero.[6] This may be attributed to the limited variation in the gender of English teachers, as 91% of them are female.

The second row indicates that female students outperform their male

[6] When we perform the regression analysis separately for seventh and ninth graders, we observe a more pronounced and statistically significant effect on English scores among seventh graders. The results are available upon request.

counterparts in Chinese and English by 0.47 and 0.51 standard deviations, respectively, even when taught by a male teacher.[7] However, there is no significant gender gap observed in math when students have a male math teacher. In the third row, we observe no statistically significant evidence that the academic performance of male students is affected by having a female teacher in any of the three subjects.

The results are consistent with those of Gong et al. (2018) and Xu and Li (2018), who also analyze data from the first wave of CEPS. Gong et al. (2018) report an estimated female teacher effect of a 0.18 standard deviation increase in test scores pooled over all three subjects, which aligns with the average of the estimates from separate analyses of math, Chinese, and English scores in our study. Interestingly, our study reveals greater benefits of student-teacher gender matching for female students compared to Xu and Li (2018), which document 0.13–0.15 standard deviation increases in math scores and 0.08–0.14 standard deviation increases in Chinese scores. This difference may stem from our application of stricter sample restriction criteria, resulting in a 35% smaller analysis sample that includes schools with more stringent enforcement of random classroom assignment, as in Gong et al. (2018).

### 4.2 Effects on Study Behavior

In this section, we investigate whether female students exhibit improved study behaviors when taught by female teachers compared to when taught by male teachers, or in comparison to male students taught by female teachers. While participation in out-of-school private education is measured separately for each subject, the variables related to class attendance, punctuality, and homework time are not subject-specific. As a result, the findings for these three variables should be interpreted with caution, as they measure the impact of having a female subject teacher on overall learning attitude and effort rather than within a particular subject. We provide estimates from regressions with controls in Table 4 and without controls in Appendix Table 2, and their similarity once again supports the validity of the random classroom assignment.

---

[7] Xu and Li (2018) also find that the test score gap between female and male students in Chinese and English is approximately 0.5–0.6 standard deviations when taught by male teachers.

Table 4. Effects of Teacher Gender on Study Behaviors

| Dependent variable: | Skip class | Late to class | Average homework time | Private out-of-school education |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| ***Panel A. Math*** | | | | |
| Female student * Female teacher | -0.030 | -0.099* | -0.026 | 0.041* |
| | (0.054) | (0.059) | (0.066) | (0.022) |
| Female student | -0.060 | -0.029 | 0.182*** | 0.007 |
| | (0.040) | (0.045) | (0.049) | (0.018) |
| Female teacher | 0.018 | 0.018 | -0.024 | -0.019 |
| | (0.049) | (0.057) | (0.075) | (0.031) |
| Adjusted $R^2$ | 0.001 | -0.001 | -0.002 | 0.205 |
| Observations | 9,175 | 9,746 | 9,721 | 9,005 |
| ***Panel B. Chinese*** | | | | |
| Female student * Female teacher | 0.061 | -0.005 | 0.024 | -0.008 |
| | (0.060) | (0.070) | (0.066) | (0.018) |
| Female student | -0.121** | -0.082 | 0.154*** | 0.014 |
| | (0.051) | (0.062) | (0.050) | (0.014) |
| Female teacher | -0.032 | 0.068 | -0.021 | 0.009 |
| | (0.075) | (0.094) | (0.093) | (0.019) |
| Adjusted $R^2$ | -0.001 | -0.002 | -0.002 | 0.143 |
| Observations | 8,815 | 9,572 | 9,042 | 9,596 |
| ***Panel C. English*** | | | | |
| Female student * Female teacher | -0.190*** | -0.197** | -0.010 | 0.031 |
| | (0.051) | (0.083) | (0.111) | (0.026) |
| Female student | 0.089** | 0.097 | 0.177* | -0.003 |
| | (0.043) | (0.078) | (0.105) | (0.025) |
| Female teacher | 0.072 | 0.029 | -0.278*** | -0.005 |
| | (0.059) | (0.093) | (0.068) | (0.022) |
| Adjusted $R^2$ | -0.001 | -0.001 | 0.000 | 0.275 |
| Observations | 8,861 | 9,619 | 9,099 | 9,644 |

*Notes*. Skip class, late to class, and average homework time are standardized to have zero mean and unit variance within school-by-grade groups. Control variables include student and teacher characteristics. See footnote 3 for a comprehensive list of control variables. We impute missing values of binary control variables with 0, and dummies for missing observations are also controlled for. Robust standard errors in parentheses are clustered at the school-by-grade level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4 provides suggestive evidence that female students' class attendance and punctuality behaviors improve when they are assigned to female subject teachers. Specifically, we observe a 0.1 standard deviation decrease in tardy behavior when female students are taught by a female math teacher (panel A), and about a 0.2 standard deviation reduction in class absence and late attendance when they are taught by a female English teacher (panel C). There is also a slight increase in the likelihood of participating in private out-of-school education on math when female students are matched with a female math teacher. However, we find no statistically significant change in any of the study behaviors due to a Chinese female teacher. The impacts on test scores and study behaviors do not precisely align across subjects, possibly because class attendance or homework behaviors are not subject specific. Due to this data limitation, we are unable to directly demonstrate whether the rise in test scores is attributed to enhancements in subject-specific study behaviors.

When analyzing the data by grade level, we observe that the effects on class attendance and punctuality behaviors are mainly driven by ninth graders. A notable exception is a 0.26 standard deviation increase in homework time among seventh-grade female students matched with a female English teacher. More detailed results are available upon request.

### 4.3 Heterogeneous Effects

In this section, we delve into whether the advantages of having a female subject teacher vary by student background characteristics. We perform heterogeneity analyses based on family financial status and parents' highest education level. Additionally, we employ quantile regression methods to examine whether test score improvements due to student-teacher gender match differ between high- and low-achieving students.

***By family financial status.*** Table 5 presents the regression results on test scores and study behaviors by family financial status. In columns (1)–(3), we find that the positive impact of female math teachers on female students' math performance is substantially larger for students with economically disadvantaged background. In poor families, female students' math test score improves by 0.315 standard deviations compared to male students, while it improves by 0.223 standard deviations in middle-income families and no significant change is observed in rich families. We do not observe such patterns for Chinese or English.

Table 5. Heterogeneous Effects on Test Scores and Study Behaviors, by Family Financial Status

| Dependent variable: | Test score | | | Skip class | | | Late to class | | | Average homework time | | | Private out-of-school education | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Poor (1) | Moderate (2) | Rich (3) | Poor (4) | Moderate (5) | Rich (6) | Poor (7) | Moderate (8) | Rich (9) | Poor (10) | Moderate (11) | Rich (12) | Poor (13) | Moderate (14) | Rich (15) |
| ***Panel A. Math*** | | | | | | | | | | | | | | | |
| Female student * Female teacher | 0.315*** | 0.223*** | 0.058 | 0.096 | -0.078 | 0.169 | 0.126 | -0.201*** | 0.017 | 0.083 | -0.008 | -0.352 | 0.053* | 0.048 | -0.064 |
| | (0.118) | (0.082) | (0.285) | (0.118) | (0.063) | (0.243) | (0.106) | (0.066) | (0.211) | (0.135) | (0.066) | (0.353) | (0.028) | (0.029) | (0.120) |
| Female student | -0.002 | -0.002 | -0.047 | -0.140** | -0.005 | -0.318* | -0.200*** | 0.082 | -0.071 | 0.232*** | 0.173*** | 0.092 | 0.007 | 0.001 | 0.115 |
| | (0.074) | (0.063) | (0.254) | (0.070) | (0.050) | (0.172) | (0.061) | (0.055) | (0.151) | (0.077) | (0.052) | (0.297) | (0.015) | (0.024) | (0.114) |
| Female teacher | 0.351 | 0.035 | 0.513** | -0.002 | 0.040 | -0.223 | -0.047 | 0.071 | -0.250 | -0.079 | -0.052 | 0.058 | -0.008 | -0.025 | -0.098 |
| | (0.306) | (0.184) | (0.248) | (0.141) | (0.062) | (0.289) | (0.099) | (0.067) | (0.226) | (0.103) | (0.084) | (0.280) | (0.031) | (0.033) | (0.162) |
| Adjusted $R^2$ | 0.033 | 0.040 | 0.108 | -0.009 | 0.000 | 0.052 | -0.014 | 0.003 | 0.129 | -0.003 | -0.004 | 0.012 | 0.144 | 0.205 | 0.120 |
| Observations | 1,825 | 7,082 | 585 | 1,805 | 6,645 | 511 | 1,855 | 7,217 | 601 | 1,740 | 6,818 | 568 | 1,867 | 7,225 | 603 |
| ***Panel B. Chinese*** | | | | | | | | | | | | | | | |
| Female student * Female teacher | 0.319** | 0.150** | 0.597 | 0.156 | 0.017 | 0.245 | 0.033 | -0.042 | 0.232 | 0.239** | -0.034 | 0.293 | 0.032 | -0.026 | 0.138 |
| | (0.154) | (0.075) | (0.378) | (0.107) | (0.076) | (0.460) | (0.103) | (0.091) | (0.333) | (0.120) | (0.073) | (0.373) | (0.025) | (0.021) | (0.162) |
| Female student | 0.472*** | 0.464*** | -0.048 | -0.219*** | -0.061 | -0.453 | -0.206** | -0.004 | -0.287 | 0.115 | 0.192*** | -0.454 | 0.004 | 0.020 | -0.057 |
| | (0.105) | (0.068) | (0.354) | (0.076) | (0.068) | (0.428) | (0.079) | (0.083) | (0.330) | (0.071) | (0.056) | (0.386) | (0.016) | (0.016) | (0.155) |
| Female teacher | -0.344 | 0.028 | -0.751 | 0.021 | -0.086 | -0.032 | 0.228* | -0.028 | -0.056 | -0.225** | -0.014 | 0.382 | -0.028 | 0.038 | -0.249* |
| | (0.282) | (0.186) | (0.575) | (0.171) | (0.073) | (0.902) | (0.126) | (0.087) | (0.546) | (0.113) | (0.118) | (0.376) | (0.021) | (0.027) | (0.145) |
| Adjusted $R^2$ | 0.131 | 0.121 | 0.139 | -0.013 | 0.000 | -0.011 | -0.009 | -0.000 | 0.105 | -0.012 | -0.002 | 0.028 | 0.072 | 0.166 | 0.135 |
| Observations | 1,783 | 6,992 | 570 | 1,764 | 6,511 | 493 | 1,815 | 7,121 | 586 | 1,706 | 6,730 | 554 | 1,827 | 7,128 | 588 |
| ***Panel C. English*** | | | | | | | | | | | | | | | |
| Female student * Female teacher | 0.036 | 0.131 | 0.076 | -0.016 | -0.258*** | -0.106 | -0.094 | -0.251* | 0.658* | 0.171 | -0.075 | -0.539 | 0.015 | 0.038 | 0.088 |
| | (0.138) | (0.108) | (0.410) | (0.145) | (0.054) | (0.334) | (0.121) | (0.128) | (0.385) | (0.178) | (0.117) | (1.290) | (0.049) | (0.032) | (0.111) |
| Female student | 0.632*** | 0.436*** | 0.309 | -0.110 | 0.185*** | -0.137 | -0.096 | 0.202 | -0.668* | 0.098 | 0.236** | 0.381 | 0.011 | -0.014 | 0.009 |
| | (0.140) | (0.097) | (0.409) | (0.147) | (0.047) | (0.297) | (0.115) | (0.126) | (0.349) | (0.164) | (0.109) | (1.276) | (0.045) | (0.030) | (0.099) |
| Female teacher | 0.273* | 0.018 | 0.250 | -0.179 | 0.198*** | -0.030 | 0.021 | 0.089 | -0.109 | -0.243** | -0.286*** | -0.205 | -0.025 | -0.004 | 0.206** |
| | (0.161) | (0.122) | (0.387) | (0.134) | (0.057) | (0.162) | (0.117) | (0.100) | (0.311) | (0.095) | (0.084) | (0.597) | (0.041) | (0.034) | (0.096) |
| Adjusted $R^2$ | 0.139 | 0.118 | 0.112 | -0.015 | 0.000 | 0.044 | -0.014 | 0.001 | 0.154 | -0.011 | 0.001 | 0.012 | 0.143 | 0.280 | 0.244 |
| Observations | 1,798 | 7,032 | 582 | 1,775 | 6,538 | 504 | 1,826 | 7,148 | 597 | 1,717 | 6,768 | 564 | 1,836 | 7,158 | 599 |

*Notes*. Test scores, skip class, late to class and average homework time are standardized to have zero mean and unit variance within school-by-grade groups. Control variables include student and teacher characteristics. See footnote 3 for a comprehensive list of control variables. We impute missing values of binary control variables with 0, and dummies for missing observations are also controlled for. Robust standard errors in parentheses are clustered at the school-by-grade level. *** $p$<0.01, ** $p$<0.05, * $p$<0.1

Regarding study behaviors, as reported in columns (4)-(15), it is notable that female students from middle-income families experience significant improvements in class attendance behaviors when taught by female teachers. Specifically, when female students from families with moderate income are assigned to a female math or English teacher, there is a substantial improvement in their overall class participation. However, female students from either poor or affluent families do not exhibit significant improvements in their class attendance behaviors. The effect of having a female Chinese teacher on study behaviors are mostly insignificant, except for the average homework time among low-income students. Overall, we observe different heterogeneous patterns by family income for test scores and study behaviors.

***By parents' education level.*** Table 6 explores the impact of teacher gender on students' academic performance and study behaviors based on the highest educational attainment of parents. We group parents' education levels in two categories: 'middle school or below' and 'high school or above,' using the same grouping as Gong et al. (2018). The heterogeneous effects on test scores by parents' education exhibit similar patterns as in the subgroup analysis by family income: disadvantaged female students experience a significantly larger test score improvement when matched with a female teacher, with this pattern being more pronounced in mathematics.

For study behaviors, on the other hand, we find no significant difference in teacher gender effects on female students by parents' education. The only exception is that participation in private out-of-school education on math significantly increases among female students whose parents have educational attainment at most up to middle school.

***Quantile regression results on test scores.*** In Appendix Table 3, we also perform quantile regression analysis for students at the $25^{th}$, $50^{th}$, and $75^{th}$ percentiles of test scores. Low-achieving female students experience a more substantial improvement in their math scores when taught by a female math teacher, with an increase of 0.376 standard deviations. This pattern is less pronounced for Chinese, and there is no evidence of heterogeneous effects across the test score distribution for English.

## Table 6. Heterogeneous Effects on Test Scores and Study Behaviors, by Parents' Education

| Dependent variable: | Test score | | Skip class | | Late for class | | Average homework time | | Private out-of-school education | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Middle school or less (1) | High school or more (2) | Middle school or less (3) | High school or more (4) | Middle school or less (5) | High school or more (6) | Middle school or less (7) | High school or more (8) | Middle school or less (9) | High school or more (10) |
| ***Panel A. Math*** | | | | | | | | | | |
| Female student * | 0.273*** | 0.139* | -0.069 | -0.042 | -0.090 | -0.193** | 0.049 | -0.100 | 0.046** | -0.032 |
| Female teacher | (0.088) | (0.077) | (0.064) | (0.091) | (0.077) | (0.089) | (0.076) | (0.093) | (0.023) | (0.034) |
| Female student | -0.007 | 0.071 | -0.077* | 0.009 | -0.053 | 0.082 | 0.157*** | 0.229*** | -0.015 | 0.092*** |
| | (0.068) | (0.065) | (0.044) | (0.076) | (0.047) | (0.078) | (0.054) | (0.082) | (0.016) | (0.032) |
| Female teacher | 0.230 | 0.019 | 0.042 | 0.065 | 0.055 | 0.032 | -0.009 | -0.091 | 0.020 | -0.049 |
| | (0.236) | (0.165) | (0.071) | (0.073) | (0.076) | (0.085) | (0.076) | (0.124) | (0.033) | (0.045) |
| Adjusted $R^2$ | 0.030 | 0.075 | -0.002 | 0.017 | -0.007 | 0.014 | -0.006 | 0.017 | 0.153 | 0.199 |
| Observations | 5,413 | 4,105 | 5,230 | 3,750 | 5,526 | 4,169 | 5,185 | 3,968 | 5,529 | 4,190 |
| ***Panel B. Chinese*** | | | | | | | | | | |
| Female student * | 0.250*** | 0.076 | 0.059 | 0.028 | -0.054 | 0.128 | 0.110 | -0.177 | -0.008 | -0.037 |
| Female teacher | (0.077) | (0.130) | (0.063) | (0.116) | (0.076) | (0.103) | (0.071) | (0.125) | (0.017) | (0.028) |
| Female student | 0.463*** | 0.485*** | -0.149*** | -0.036 | -0.068 | -0.149 | 0.111** | 0.309*** | 0.014 | 0.038* |
| | (0.061) | (0.123) | (0.050) | (0.106) | (0.061) | (0.098) | (0.055) | (0.117) | (0.012) | (0.022) |
| Female teacher | -0.162 | -0.017 | -0.023 | -0.079 | 0.117 | -0.226* | -0.125 | 0.223 | 0.011 | 0.033 |
| | (0.228) | (0.183) | (0.083) | (0.167) | (0.100) | (0.121) | (0.107) | (0.160) | (0.020) | (0.035) |
| Adjusted $R^2$ | 0.114 | 0.153 | -0.003 | 0.011 | -0.007 | 0.008 | -0.006 | 0.016 | 0.071 | 0.185 |
| Observations | 5,302 | 4,071 | 5,103 | 3,687 | 5,410 | 4,135 | 5,080 | 3,938 | 5,412 | 4,156 |
| ***Panel C. English*** | | | | | | | | | | |
| Female student * | 0.152 | 0.007 | -0.214*** | -0.157 | -0.199** | -0.274 | -0.004 | 0.127 | 0.022 | 0.017 |
| Female teacher | (0.097) | (0.089) | (0.057) | (0.240) | (0.090) | (0.175) | (0.115) | (0.180) | (0.025) | (0.039) |
| Female student | 0.486*** | 0.508*** | 0.077* | 0.136 | 0.081 | 0.224 | 0.176* | 0.040 | -0.005 | 0.024 |
| | (0.100) | (0.078) | (0.041) | (0.237) | (0.082) | (0.170) | (0.106) | (0.173) | (0.023) | (0.036) |
| Female teacher | 0.049 | 0.165 | 0.033 | 0.103 | 0.022 | 0.109 | -0.254*** | -0.442** | 0.006 | -0.016 |
| | (0.124) | (0.231) | (0.066) | (0.186) | (0.093) | (0.174) | (0.092) | (0.177) | (0.019) | (0.058) |
| Adjusted $R^2$ | 0.118 | 0.151 | -0.002 | 0.012 | -0.007 | 0.009 | -0.005 | 0.020 | 0.171 | 0.276 |
| Observations | 5,357 | 4,079 | 5,149 | 3,687 | 5,456 | 4,136 | 5,138 | 3,938 | 5,460 | 4,156 |

*Notes*. Test scores, skip class, late to class and average homework time are standardized to have zero mean and unit variance within school-by-grade groups. Control variables include student and teacher characteristics. See footnote 3 for a comprehensive list of control variables. We impute missing values of binary control variables with 0, and dummies for missing observations are also controlled for. Robust standard errors in parentheses are clustered at the school-by-grade level. *** $p$<0.01, ** $p$<0.05, * $p$<0.1

# 5 Conclusion

In this study, we estimate the effects of female teachers on test scores and study behaviors of female seventh and ninth graders using data from the China Education Panel Survey. We exploit random assignment of students to classrooms and teachers in Chinese middle schools. Our analysis shows that female teachers significantly improve female students' academic performance in their subject, particularly in math and Chinese, with test score gains of around 0.2 standard deviations. Additionally, female students' class attendance improves, with a reduction in class absence and tardiness of 0.1–0.2 standard deviations when taught by a female teacher. While the positive effect on math scores is more pronounced among disadvantaged or low-achieving students, this pattern is not evident in test scores for other subjects or in study behaviors.

The results of this study have three implications. First, the results from heterogeneity and distributional analyses suggest that pairing female students with female instructors can be an effective policy tool to enhance girls' performance and interests in STEM fields, especially among the disadvantaged or low-achieving populations. Second, the improvement in class attendance or homework behaviors may not be the primary mechanism mediating the benefit of student-teacher gender matching on test score increase, as we find inconsistent heterogeneous effects by socioeconomic status for test scores and study behaviors. Therefore, precise mechanisms require further investigation in future research. Lastly, while our study focuses on short-term outcomes within three years, exploring the duration of improvements in academic outcomes and their potential impact on career perspectives and labor market performance would be a promising avenue for future research.

## References


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

Antecol, H., Eren, O. and S. Ozbeklik, "The effect of teacher gender on student achievement in primary school," *Journal of Labor Economics* 33(1), 2015, 63-89.

Card, D., Domnisoru, C., Sanders, S.G., Taylor, L. and V. Udalova, "The impact of female teachers on female students' lifetime well-being," NBER Working paper 30430, National Bureau of Economic Research: Cambridge Massachusetts, 2022.

Carrell, S.E., Page, M.E. and J.E. West, "Sex and science: How professor gender perpetuates the gender gap," *The Quarterly Journal of Economics* 125(3), 2010, 1101-1144.

Eble, A. and F. Hu, "Child beliefs, societal beliefs, and teacher-student identity match," *Economics of Education Review* 77, 2020, 101994.

Gong, J., Lu, Y. and H. Song, "The effect of teacher gender on students' academic and noncognitive outcomes," *Journal of Labor Economics* 36(3), 2018, 743-778.

Lim, J. and J Meer, "The impact of teacher-student gender matches: random assignment evidence from South Korea," *Journal of Human Resources* 52(4), 2017, 979-997.

Lim, J. and J. Meer, "Persistent effects of teacher-student gender matches," *Journal of Human Resources* 55(3), 2020, 809-835.

Oster, E., "Unobservable selection and coefficient stability: theory and evidence," *Journal of Business & Economic Statistics* 37(2), 2019, 187-204.

Pei, Z., Pischke, J. and H. Schwandt, "Poorly measured confounders are more useful on the left than on the right," *Journal of Business & Economic Statistics* 37(2), 2019, 205-216.

Xu, D. and Q. Li, "Gender achievement gaps among Chinese middle school students and the role of teachers' gender," *Economics of Education Review* 67, 2018, 82-93.

## Appendix Table 1. List of Key Variables

| Variable | Description |
|---|---|
| Test scores | Student's midterm exam scores for the fall semester of 2013, on three major subjects of math, Chinese and English. The criteria for assessing test score are provided directly by the school through surveys. |
| Skip class | Students are asked about "Do you agree that you frequently Skip class?". The original data are tallied using a 4-point Likert scale (1 = completely disagree, 4 = completely agree). |
| Late to class | Students are asked about "Do you agree that you are often late?". The original data are tallied using a 4-point Likert scale (1 = completely disagree, 4 = completely agree). |
| Average homework time | The average time devoted to doing homework per day over the course of a week, using the date when they completed the questionnaire as a reference point. |
| Private out-of-school education | Private out-of-school education by subject: tutoring in general Chinese courses and writing; Mathematics Olympiad and general math courses; and general English courses. |
| Co-residence with parents | Whether the student live with both parents. |
| Only child | Whether the student is the only child. |
| Father often gets drunk | Whether the student's father often gets drunk. |
| Father and mother often quarrel | Whether the student's father and mother often quarrel. |
| Parents have a good relationship | Whether the student's parents are in a good relationship. |
| Graduated from 4-year college | Whether the teacher holds a bachelor's degree from a 4-year college or higher. |
| Have over 20 years of teaching experience | Whether the teacher has an accumulation of 20 or more years of teaching experience. |
| Major in education in college | Whether the teacher graduated from an education college or pursued education as a major at a non-education college. |
| Have administrative positions | Whether the teacher holds administrative positions, which include head of the grade, head of the teaching and research group of a subject, teaching director or deputy teaching director, and deputy principal. |

*Notes*. Test scores, skip class, late to class and average homework time are standardized to have zero mean and unit variance within school-by-grade groups.

Appendix Table 2. Effects of Teacher Gender on Study Behaviors, without Controls

| Dependent variable: | Skip class | Late to class | Average homework time | Private out-of-school education |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| ***Panel A. Math*** | | | | |
| Female student * Female teacher | -0.036 | -0.107* | -0.027 | 0.043** |
| | (0.053) | (0.058) | (0.067) | (0.022) |
| Female student | -0.070* | -0.034 | 0.186*** | 0.005 |
| | (0.039) | (0.045) | (0.051) | (0.017) |
| Female teacher | -0.010 | -0.007 | -0.004 | -0.001 |
| | (0.053) | (0.061) | (0.055) | (0.027) |
| Adjusted $R^2$ | -0.010 | -0.010 | -0.006 | 0.201 |
| Observations | 9,005 | 9,721 | 9,175 | 9,746 |
| ***Panel B. Chinese*** | | | | |
| Female student * Female teacher | 0.061 | -0.003 | 0.034 | -0.011 |
| | (0.062) | (0.070) | (0.067) | (0.017) |
| Female student | -0.130** | -0.089 | 0.151*** | 0.013 |
| | (0.054) | (0.061) | (0.053) | (0.014) |
| Female teacher | 0.004 | 0.117 | -0.014 | 0.007 |
| | (0.069) | (0.098) | (0.106) | (0.019) |
| Adjusted $R^2$ | -0.010 | -0.010 | -0.006 | 0.138 |
| Observations | 8,815 | 9,572 | 9,042 | 9,596 |
| ***Panel C. English*** | | | | |
| Female student * Female teacher | -0.206*** | -0.219*** | -0.008 | 0.030 |
| | (0.045) | (0.082) | (0.110) | (0.027) |
| Female student | 0.094*** | 0.108 | 0.177* | -0.004 |
| | (0.035) | (0.077) | (0.104) | (0.026) |
| Female teacher | 0.098* | 0.045 | -0.288*** | -0.010 |
| | (0.059) | (0.104) | (0.065) | (0.024) |
| Adjusted $R^2$ | -0.009 | -0.009 | -0.002 | 0.271 |
| Observations | 8,861 | 9,619 | 9,099 | 9,644 |

*Notes*. Skip class, late to class and average homework time are standardized to have zero mean and unit variance within school-by-grade groups. Robust standard errors in parentheses are clustered at the school-by-grade level. *** $p$<0.01, ** $p$<0.05, * $p$<0.1

## Appendix Table 3. Distributional Effects of Teacher Gender on Test Scores

| **Quantiles:** | **0.25** | **0.5** | **0.75** |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***Panel A. Math*** | | | |
| Female student * Female teacher | 0.376*** | 0.190*** | 0.078* |
| | (0.062) | (0.056) | (0.047) |
| Female student | 0.038 | 0.032 | 0.000 |
| | (0.048) | (0.045) | (0.037) |
| Female teacher | 0.253*** | 0.208*** | 0.128*** |
| | (0.059) | (0.058) | (0.043) |
| Observations | 9,548 | 9,548 | 9,548 |
| ***Panel B. Chinese*** | | | |
| Female student * Female teacher | 0.183** | 0.184*** | 0.100* |
| | (0.071) | (0.059) | (0.052) |
| Female student | 0.646*** | 0.422*** | 0.356*** |
| | (0.059) | (0.050) | (0.048) |
| Female teacher | -0.450*** | -0.268*** | -0.100 |
| | (0.071) | (0.081) | (0.079) |
| Observations | 9,404 | 9,404 | 9,404 |
| ***Panel C. English*** | | | |
| Female student * Female teacher | 0.023 | 0.106 | -0.043 |
| | (0.110) | (0.106) | (0.078) |
| Female student | 0.712*** | 0.527*** | 0.412*** |
| | (0.105) | (0.102) | (0.075) |
| Female teacher | -0.117 | -0.245*** | 0.080 |
| | (0.085) | (0.094) | (0.066) |
| Observations | 9,467 | 9,467 | 9,467 |

*Notes*. Test scores are standardized to have zero mean and unit variance within school-by-grade groups. Regressions do not control for student and teacher characteristics. Robust standard errors in parentheses. *** $p$<0.01, ** $p$<0.05, * $p$<0.1