

Bridging Education Gender Gaps in Developing Countries: The Role of Female Teachers

Karthik Muralidharan

Ketki Sheth

We study gender gaps in learning, and the effectiveness of female teachers in reducing them, using a large, representative, annual panel dataset from the Indian state of Andhra Pradesh. We find a small but significant negative trend in girls' test-scores in both math and language. Using five years of panel data, we find that teachers are more effective at teaching students of their own gender. Female teachers are more effective at teaching girls than male teachers but no worse at teaching boys. Thus, hiring female teachers on the current margin may reduce gender gaps in test scores without hurting boys.

Karthik Muralidharan is an associate professor in the Department of Economics at the University of California, San Diego. Ketki Sheth is an assistant professor in economics at the University of California, Merced. The authors thank Prashant Bharadwaj, Julie Cullen, Gordon Dahl, Craig McIntosh, Miguel Urquiola, two anonymous referees, and several seminar participants for comments. The authors also thank the AP RESt team and the Azim Premji Foundation for collecting the data used in this paper, and thank Venkatesh Sundararaman for the overall support provided to the AP RESt project. Financial assistance for the data collection was provided by the Government of Andhra Pradesh, the UK Department for International Development (DFID), the Azim Premji Foundation, and the World Bank. The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of any of the organizations that supported the data collection. The data used in this article can be obtained beginning six months after publication through three years hence from Karthik Muralidharan, 9500 Gilman Drive, San Diego, CA, kamurali@ucsd.edu.

I. Introduction

Reducing gender gaps in education attainment has been an important priority for international education policy and is explicitly listed as one of the United Nations Millennium Development Goals (MDGs). This commitment has been reflected in the policies of many developing countries, and substantial progress has been made in the past decade in reducing gender barriers in primary school enrollment. One key policy that is credited with increasing girls' education is the increased recruitment of female teachers (UNESCO 2012; Herz and Sperling 2004). UNICEF has documented the practice in a variety of countries, including Bangladesh, India, Liberia, Nepal, and Yemen, and the United Nations' Task Force for achieving the MDGs has advocated hiring more female teachers as an effective policy mechanism for reaching the goal of universal primary education of girls (UNDG 2010; Rehman 2008; Slavin 2006).

While the idea that hiring more female teachers can bridge gender gaps is widely prevalent among policy makers, there is very little empirical evidence from testing this hypothesis in developing countries. In this paper, we study the causal impact of having a female teacher on the learning gains of female students using one of the richest datasets on primary education in a developing country. The dataset features annual longitudinal data on student learning measured through independent assessments conducted over five years across a representative sample of 500 rural schools and over 90,000 students in the Indian state of Andhra Pradesh (AP). The data also includes detailed information on teacher characteristics and on their assignments to specific classrooms in each year.

The combination of panel data and variation in the gender of teachers and students allows us to estimate the causal impact of matching teacher and student gender in a value-added

framework. Identification concerns are addressed by showing that our estimates of gender matching do not change under an increasingly restrictive set of specifications including school, school-grade, teacher, and student gender by grade fixed effects. We also show that there is no correlation between the probability of being assigned a female teacher and either the fraction of female students in the class or the mean test scores at the start of the year. Further, our estimation sample is restricted to schools that only have one section per grade, which precludes the possibility that students may be tracked across sections and that female teachers may be assigned to different sections based on unobservables.

We report five main findings in this paper. First, we find a small but significant negative trend in girls' test scores in both math ($0.02\sigma/\text{year}$) and language ($0.01\sigma/\text{year}$) as they advance through the five grades of primary school.¹ Girls have significantly higher test scores in language and equal test scores in math relative to boys at the end of first grade, but score almost on par with boys in language and significantly worse in math by the end of grade five. These results are consistent with evidence of gender gaps in test scores (particularly in math) documented in both high and low income countries (Fryer and Levitt 2010; Bharadwaj et al. 2012), and suggest that the growing gender gaps documented at later ages in both these papers probably reflect a cumulative effect of a trend that starts as early as primary school.

Second, using five years of panel data and school-grade and student gender by grade fixed effects, we find that teachers are $.034\sigma/\text{year}$ more effective in teaching students of their own gender *relative* to teachers of the opposite gender. In other words, female teachers are $.034\sigma/\text{year}$ more effective at reducing the gender gap in achievement than male teachers. Since female teachers differ from male teachers on several characteristics that may be correlated with teacher quality, we test the robustness of the 'gender-match' result by including interactions

between student gender and each of the teacher characteristics on which female and male teachers differ and find that our estimates are essentially unchanged.

The result above is a difference-in-difference estimate that compares the relative advantage of female teachers in teaching girls rather than boys with the relative disadvantage of male teachers in teaching girls rather than boys. However, the overall effectiveness of a teacher is also determined by his or her effectiveness at teaching students of the opposite gender. Our third result speaks to this issue, and we find that female teachers in our setting are more effective overall than male teachers. We find that girls who have a female teacher in a given year have $.036\sigma$ higher annual test score gains than if she had a male teacher. However, boys perform similarly regardless of the gender of their teacher. Thus, girls are likely to benefit from a policy of hiring more female teachers, and overall educational performance is likely to increase due to the lack of any offsetting effect on boys.

Fourth, we study the impacts of a teacher-student gender match on student attendance and find no evidence that teachers are more effective at raising the attendance for students of the same gender. This suggests that the likely mechanism for the 'matching' effect on test scores is not on the extensive margin of increased student-teacher contact time, but rather on the intensive margin of more effective classroom interactions.

Finally, we document that female teachers are more likely to teach in earlier grades. Combined with the results above, we estimate that around 10-20 percent of the trend of increasing gender gaps in test scores over time can be attributed to the reduction in the probability of girls being taught by female teachers as they advance to higher grades. Since teachers in higher grades are more likely to be male across several countries (UNESCO 2010),

our results suggest that one possible channel for growing gender gaps in achievement (especially in math) could be the reduced likelihood of having female teachers in higher grades.

Our results suggest that the causal impact of having a female teacher (relative to a male teacher) on the annual learning gains of girls is positive in this setting, with no adverse impact on the learning gains of boys. Since controlling for observable teacher characteristics such as education, training, experience, salary, contractual status, union membership, place of origin, and absence rates (and their interaction with student gender) does not change these results, it *must* be the case that there are other unobservable differences across teacher gender that are driving our results. If we were able to identify an observable teacher characteristic that made the 'female teacher' effect insignificant, it would be possible to target hiring on that characteristic in lieu of gender. However, since we are not able to identify such an observable mechanism for our estimated effect, a useful way of interpreting our results is that teacher gender may be a summary statistic for unobserved teacher characteristics (such as empathy, classroom management skills, or role model effects) that are not used in hiring decisions under the status quo.

While there have been several studies on the impact of shared gender between teachers and students on learning outcomes in high income country contexts, there is surprisingly little well-identified evidence on this question in developing countries. In the US and UK, studies have shown improved test scores, teacher perception, student performance, and engagement of girls when taught by a female teacher in schools, with magnitudes of test score impacts similar to those found in our paper (Dee 2007; Dee 2005; Nixon and Robinson 1999; Ehrenberg, Goldhaber, and Brewer 1995; Ouazad and Page 2012). However, other studies conducted in both the US and in European countries have failed to find such an effect (Holmund and Sund 2008; Carrington, Tymms, and Merrell 2008; Lahelma 2000; Winters et al. 2013; Marsh, Martin, and

Chend 2008; Driessen 2007; Neugebauer, Helbig, and Landmann 2011), and some research even suggests that female teachers may adversely affect girls' performance in areas where girls face larger stereotypes (Antecol, Eren, and Ozbeklik 2015). In higher education institutions in the US, female professors have been found to have small effects on female students' course selection, achievement, and major choice (Bettinger and Long 2005; Carrell, Page, and West 2010; Hoffmann and Oreopoulos 2009).²

The question of the role of female teachers in reducing gender gaps is much more salient in low income country contexts, where gender gaps in school enrollment and attainment are much larger (OECD 2010; Hausmann, Tyson, and Zahidi 2012; Muralidharan and Prakash 2013; Bharadwaj et al. 2012), and where increased recruitment of female teachers has been actively advocated (UNDG 2010). The only related paper in a developing country setting is Rawal and Kingdon (2010), who use test score data on second and fourth grade students in the Indian states of Bihar and Uttar Pradesh and find a positive impact on educational achievement for girls taught by female teachers, but find no similar effect for boys. Since, the literature from the US and Europe may not be transferrable to developing countries (given the larger prevalence of gender stereotypes and gender gaps in these settings), our estimates fill an important gap by providing among the first estimates of the impact of teacher-student gender matching in a developing country.

In addition to providing well-identified estimates of the impact of matching teacher and student gender on learning outcomes in a developing country, where the literature is very sparse, our dataset allows us to make advances relative to both the developed and developing country literatures on this subject. First, while several existing papers in this literature (especially those looking at college-level outcomes) use grades or test scores assigned by the students' own

teachers, the test scores used in this paper are based on independent assessments and grading. This allows us to be confident that the effects we measure reflect genuine impacts on learning by eliminating the concern that the measured effects of gender matching may reflect more generous grading by teachers towards students who share their own gender.

Second and more important, the majority of papers in the global literature on this question (including Dee 2007 and Rawal and Kingdon 2010) use student fixed effects and variation in the gender of teachers across different subjects to identify the impact of the gender match on learning, but they are based on comparing *levels* of test scores as opposed to *value-added*. Thus, it can be difficult to interpret the magnitudes of the estimated effects without knowing the gender composition of the teachers in that subject in previous grades.³ Our use of five years of annual panel data on test scores allows us to estimate the impact of a gender match on the *value-added in the year that the match occurred*, which has a much clearer interpretation relative to the standard in the literature.

Finally, we observe students at a younger and more formative age than most of the literature, when the role of sharing gender with teachers may be especially important. This is also the age that is most relevant to policy for reducing education gender gaps in developing countries since the majority of students do not complete more than eight years of school education. Our estimates, based on a large dataset that is representative of the rural public school system in a state with over 80 million people, are also likely to have more external validity across other developing countries than existing work.

II. Context and Dataset

India has the largest primary schooling system in the world, catering to over 200 million children. As in other developing countries, education policy in India has placed a priority on

reducing gender disparities in education, and both the Five Year Plans and Sarva Shiksha Abhiyan (SSA), the flagship national program for universal primary education, have called for an increase in recruiting female teachers as a policy for increasing girls' education. SSA requires that 50 percent of new teachers recruited be women, and the 11th Five Year Plan suggested that it be increased to 75 percent (Government of India 2008). These calls for increased female teachers reflect a belief that through such mechanisms as role model effects, increased safety, reduced prejudices, and greater identification and empathy, female teachers are arguably more effective in increasing girls' achievement in primary school relative to their male counterparts (Ehrenberg, Goldhaber, and Brewer 1995; Stacki 2002; Dee 2005).

This paper uses data from the Indian state of Andhra Pradesh (AP), which is the 5th most populous state in India, with a population of over 80 million (70 percent rural).⁴ The data was collected as part of the Andhra Pradesh Randomized Evaluation Studies (AP RESt), a series of experimental studies designed to evaluate the impact of various input and incentive-based interventions on improving education outcomes in AP.⁵ The project collected detailed panel data over five years (covering the school years 2005-06 to 2009-10) on students, teachers, and households in a representative sample of 500 government-run primary schools (grades one through five) across five districts in AP. The dataset includes annual student learning outcomes as measured by independently conducted and graded tests in language (Telugu) and math (conducted initially at the start of the 2005-06 school year as a baseline, and subsequently at the end of each school year), basic data on student and teacher demographics, and household socioeconomic data for a subset of households. The assessments were created (based on the pedagogical objectives of the curriculum), administered, and externally graded by an independent agency, ensuring that the tests are valid measures of learning and that the scores are

not biased by teacher subjectivity. The test scores are normalized within each year-grade-subject combination and all analysis is conducted in terms of normalized test scores, with magnitudes being reported in standard deviations.

The Appendix⁶ provide further details on the dataset, including sample size by cohort (Table A1), testing for changes in characteristics of incoming cohorts by gender over time (Table A2), and differential attrition by gender (Table A3 – Panel A). There is some differential attrition in the sample over time by gender (where attrition is defined as the fraction of students who had taken a test at the end of year 'n-1', but did not take a test at the end of year 'n'), with male students more likely to attrite (around 3 percent each year). However, this differential attrition is not a first-order concern for estimating the impact of a gender-match on test score gains because we see that there is no differential attrition by student gender as a function of gender-match during the school year (Table A3 – Panel B).⁷

Table 1 – Panel A presents descriptive statistics on students who have at least one recorded test score and data on gender in the dataset.⁸ Girls comprise 51 percent of the sample of public-school students in our sample. This does not imply that more girls are going to school than boys as it is likely that more boys are attending private schools (Pratham 2012). However, it does illustrate that on average, girls are well represented in public primary schools. The girls in the sample come from modestly better off socioeconomic backgrounds than the boys and have parents who are slightly more educated and affluent. These differences probably reflect two dimensions of selection into the sample – better off households are more likely to send girls to school, and better off households are more likely to send boys to private schools. However, the magnitudes of these differences are quite small (often less than two percentage points), and the statistical significance reflects the very large sample size. Since the household surveys were

completed for only 70 percent of the sample of students for whom we have test score data, our main specifications do not include household controls.⁹

Table 1 – Panel B presents summary statistics for the teachers in our analysis. Female teachers comprise 46 percent of the total teacher body, but are less experienced, less likely to have completed high school or a masters degree, and less likely to hold a head-teacher position. Not surprisingly, their mean salaries are also lower. They also comprise a much greater share of the contract teacher workforce than that of regular civil-service teachers. Since teacher characteristics vary systematically by gender, we will report our key results on the impact of matching teacher and student gender both with and without controls for these additional teacher characteristics. We will also examine the extent to which our main results on the effects of a teacher-student 'gender match' on learning outcomes can be explained by these observable differences in teacher characteristics by gender by including interactions of student gender with each of the teacher characteristics that are different across male and female teachers.

Table 2 – Panel A presents summary statistics on gender differences in test scores by grade. We see that girls score as well as boys in math and score 0.05σ *higher* on language in grade one. However, there is a steady decline in girls' test scores in both math and language as they advance to higher grades, and by the last two years of primary school (grades four and five) we see that girls' initial advantage in language scores has declined and they do significantly worse than boys in math (by around 0.1σ). Table 2 – Panel B quantifies the annual decline in girls' relative scores by including an interaction term between student gender and grade in a standard value-added specification. We find evidence of a growing education gender gap among test takers in public primary schools, with a mean decline of 0.02σ /year in math scores and 0.01σ /year in language

scores for girls relative to boys. The results are also robust to including school fixed effects, suggesting that these differential trends are present both across and within schools.

One caveat to the interpretation of the above numbers is that they are based on a representative sample of test-taking students in public schools. Relative to the gender gap in the universe of primary-age school children, our estimate may be biased downwards if higher-scoring boys are differentially more likely to leave public school to attend private schools. Conversely, they may be biased upwards if lower-scoring boys are more likely to be absent on the day of testing. We see some evidence of the second concern because girls who are absent from the test have slightly higher previous test scores compared to boys absent from the test (Table A3 – Panel B).¹⁰

We address this concern by reweighting the estimates in Table 2 – Panel B to account for the differential probability of attrition by gender at each value of lagged test scores (we assign each observation a weight that is equal to the inverse of its probability of remaining in the sample). This reweighting is analogous to simulating the missing scores of attritors using their lagged test scores, which is the procedure followed by Bharadwaj, Loken, and Nielson (2013). We present these results in Table 2 – Panel C and see that the results from Table 2 – Panel B are unchanged, suggesting that any bias from differential attrition by gender and lagged test scores is second order (which is not surprising given the very small magnitude of test score differences among attritors by gender seen in Table A3 – Panel B). Of course, even the reweighting only provides us with estimates of gender gaps for the population of students who enter public primary schools (and cannot account for the population of students in private schools), and thus our estimates should be interpreted as relevant for the population of students in public primary schools.¹¹

III. Estimation and Identification

Our main estimating equation takes the form:

$$(1) \quad E_{itjk} = \alpha + \gamma E_{it-1j-1k} + \beta_1 F_{itjk} * g_i + \beta_2 g_i + \beta_3 F_{itjk} + \mu_{itjk}$$

where E_{itjk} are student educational outcomes (test scores and attendance) for student i , in year t , grade j , and school k respectively. F_{itjk} is an indicator for whether the student's current teacher is female, g_i is an indicator for whether the student is a girl, $F_{itjk} * g_i$ is an indicator for whether a girl student shares her teacher's gender in the current year, and μ_{itjk} is a stochastic error term.

The inclusion of the lagged test score on the right-hand side of Eq (1) allows us to estimate the impact of contemporaneous inputs in a standard value-added framework. Since all test scores are normalized by grade and subject, the estimated coefficients can be directly interpreted as the correlation between the covariate and annual gains in normalized test scores.¹² When studying attendance we do not include the lagged attendance of the previous year. We later augment Eq (1) with T_{itjk} , a vector of additional teacher characteristics, to estimate the robustness of our effects to holding other teacher characteristics constant.

The above estimating equation allows us to calculate the marginal impact of changing each component of the feasible student-teacher gender combinations relative to boys taught by male teachers (the omitted category).

The first coefficient of interest in this paper is β_1 , which indicates the extent to which teachers are relatively more effective at teaching to their own gender compared to teachers of the opposite gender. Since the indicator variable is based on the interaction of teacher and student gender, the coefficient is a 'difference-in-difference' estimate of the impact of female teachers when teaching girls rather than boys *relative* to their male counterparts teaching girls rather than boys. The coefficient on the interaction term therefore reflects the sum of the relative advantage of female teachers when teaching girls (rather than boys) and the relative disadvantage of male

teachers when teaching girls (rather than boys); specifically, $\beta_1 = (\text{female teachers teaching girls} - \text{female teachers teaching boys}) - (\text{male teacher teaching girls} - \text{male teachers teaching boys})$.

A more intuitive interpretation is to note that β_1 represents the relative effectiveness of female teachers (compared to male teachers) in reducing the test score gap between girls and boys. By construction, this is symmetric and equivalent to the relative effectiveness of male teachers teaching boys compared to girls *relative* to female teachers teaching boys compared to girls. It is important to highlight that a positive β_1 does not necessarily imply that both boys and girls have better outcomes when sharing their teacher's gender. For example, a positive β_1 could co-exist with a situation where all students are better off with female (or male) teachers because the general effectiveness of female (or male) is considerably higher (even for students of the opposite gender).

β_2 is the difference in test score gains of girls taught by male teachers relative to boys taught by male teachers; specifically, $\beta_2 = (\text{male teachers teaching girls} - \text{male teachers teaching boys})$. β_3 is the difference in test score gains of boys taught by female teachers relative to when taught by male teachers; specifically, $\beta_3 = (\text{female teachers teaching boys} - \text{male teachers teaching boys})$. Thus, β_3 estimates the extent to which boys perform differently when they are taught by a female teacher relative to a male teacher.

Starting with the omitted category (of male teachers teaching boys), adding combinations of β_1, β_2 , and β_3 allow us to measure other marginal effects of interest. Analogous to β_3 for boys, testing whether $\beta_1 + \beta_3 > 0$ provides a formal test of whether girls gain by being paired with female teachers relative to male teachers. The derivation is below:

$$(2) \text{ Female teachers teaching girls} - \text{Male teachers teaching girls} > 0$$

$$\Rightarrow (\alpha + \beta_1 + \beta_2 + \beta_3) - (\alpha + \beta_2) > 0 \Rightarrow \beta_1 + \beta_3 > 0$$

As highlighted earlier, it is possible that female teachers are relatively more effective at teaching girls than boys compared to male teachers (a positive β_1), but that female teachers are overall less effective (a negative β_3), resulting in girls being better off with male teachers despite the loss in gains from not sharing their teacher's gender ($\beta_1 + \beta_3 < 0$).

Additionally, if we value both boys' and girls' educational achievement equally, then we would be interested in knowing whether the positive gain for girls taught by female teachers outweighs any adverse effects from mismatching boys to being taught by female teachers (*potential gain to girls + potential loss to boys*). The formal test for this is $\lambda_g \beta_1 + \beta_3 > 0$, where λ_g is the proportion of girls in schools. The derivation is below:

$$(3) \lambda_g * \text{potential gain to girls} + (1 - \lambda_g) * \text{potential loss to boys} > 0$$

$$\Rightarrow \lambda_g(\beta_1 + \beta_3) + (1 - \lambda_g)(\beta_3) > 0 \Rightarrow \lambda_g \beta_1 + \beta_3 > 0$$

Thus, if the effect of female teachers on boys is negative, but their effect on girls is positive, we would find that $\beta_3 < 0$ and $\beta_1 + \beta_3 > 0$. The test outlined in Eq (3) can also be interpreted as the overall effectiveness of female teachers relative to male teachers. Intuitively, the impact of replacing a male teacher in a classroom with a female teacher is equal to the sum of the impact of the female teacher on all students (β_3), and the additional gains to female students from matching with a female teacher (β_1), weighted by the fraction of female students in the classroom (λ_g).

A. Threats to Identification

The main identification challenge in interpreting these coefficients causally is that teachers are not randomly assigned to schools, and it is possible that schools with more female teachers have greater female enrollment and are in areas that value education more and thus have steeper learning trajectories. Thus, it is possible that girls would perform well in these schools regardless

of their teacher's gender. In such a case, the estimate of β_1 could be confounded by omitted variables correlated with both the probability of having a female teacher and steeper learning trajectories for girls. We address this concern by augmenting Eq (1) with school fixed effects, and thereby estimating the impact of a gender-match on value-added *relative* to the schools' average effectiveness at improving value-added.

A further concern could be that teachers are not assigned randomly to grades within schools, and a similar omitted variable concern would apply if female teachers are differentially assigned to grades within schools in which students have higher learning trajectories and there is higher girls' enrollment. To address this concern, we include school-grade fixed effects, which controls for the average performance in a given *grade* in the school (instead of the overall performance of the school).

Finally, it could be the case that female teachers are generally assigned to grades where girls have steeper learning trajectories relative to boys. To account for such differential trajectories of learning in different grades *by student gender*, we also include student gender by grade fixed effects to estimate the parameters of interest by comparing test score gains relative to girls' and boys' *average learning trajectories in each grade*. Our preferred specification therefore includes both school-grade fixed effects and student gender by grade fixed effects to address these concerns.¹³

A final concern is that if grades in a school have multiple sections, then the assignment of teachers to sections within grades could be based on omitted variables, such as a greater probability of assigning female teachers to sections with girls who have a steeper learning trajectory. However, this is not an important factor in our setting because schools typically have fewer teachers than grades, and the typical teaching arrangement is one of multi-grade teaching

(where the same teacher simultaneously teaches multiple grades). As a result, there are only few cases where there are multiple sections per grade with different teachers assigned to different sections. We drop all such cases (6 percent of observations) where there are multiple teachers per grade. Since students are taught math and language by the same teacher in a given year for all our observations, we are unable to use student fixed effects to identify the impact of a 'gender match' using variation in the gender of teachers across different subjects.¹⁴ Note that our identification strategy does not require teacher gender to switch in a given school-grade over time, and neither does it require teacher gender to switch within a cohort over time (across different grades).¹⁵ Rather, the inclusion of school-grade and student gender by grade fixed effects implies that the identifying variation is coming from the differential effectiveness of teachers (by gender) at teaching girls versus boys relative to (a) the mean value-added experienced by students in that school and grade over the five years of data, and (b) the mean value-added for girls relative to boys in that grade across all schools in the sample.

B. Testing the Identifying Assumptions

Table 3 shows the correlation between various classroom characteristics and the probability of the classroom having a female teacher. We see that there is no significant correlation between having a female teacher and the fraction of girls in the classroom, or with the average test scores of incoming cohorts for either gender. However, female teachers *are* more likely to be assigned to younger grades. This is why our preferred specifications include school-grade fixed effects. Upon the inclusion of school-grade fixed effects, it continues to be case that there is no significant correlation between having a female teacher in the class and either the fraction of female students or the test scores of the incoming cohort (columns 5 and 6).

However, we see in Table 4 that girls do have a slightly more concave learning trajectory than boys. We estimate a standard value-added model that controls for lagged test scores (as in Eq 1), but allows for an interaction between student gender and grade, and find that female students have lower value-added in higher grades. Since female teachers are more likely to be assigned to lower grades, the inclusion of school-grade fixed effects alone (the average test score gain in a grade within a school over the five years across both student genders) does not address the possible spurious correlation from female teachers being more likely to be assigned to grades where female students fall behind boys at a lower rate. Therefore, the inclusion of grade fixed effects by student gender in our main specifications is necessary to control for average value-added test scores in each grade *by student gender*. Thus, the parameters of interest in Eq (1) are identified relative to the average learning trajectory for girls in the same grade (student gender by grade fixed effects) and relative to the average learning trajectory in the same school for that grade (school-grade fixed effects).

IV. Results

A. Test Score Impacts of Matching Teacher and Student Gender

The main results of the paper (Eq 1) are presented in Table 5, which pools student test scores across subjects (results separated by subject are in Table 7). The columns show increasingly restrictive identification assumptions with school fixed effects (Column 2), school-grade fixed effects (Column 3), and both of these with student gender by grade fixed effects (Column 4 and 5). Column 6 expands the preferred specification in Column 5 to include teacher covariates to test the extent to which *average* female teacher effects can be explained by other observable characteristics that can be used in teacher selection. Thus, the estimates in column 5 are relevant to the policy question: "What will happen if we replace a male teacher with a female teacher

whose characteristics are the same as those of the average female teacher?" On the other hand, the estimates in column 6 answer the question: "What will happen if we just switch a teacher's gender from male to female holding other commonly observed teacher recruitment characteristics constant?"

Of course, switching the gender in this latter case does still include all unobservable characteristics correlated with being a male versus female teacher. Though we cannot separately identify which of these unobservable characteristics are driving differences in average teacher effectiveness, the differences between columns 5 and 6 helps to clarify whether the 'female' teacher effect in column 5 can be explained by observable differences across male and female teachers, or whether it represents unobservable characteristics of teacher effectiveness that current recruitment practices do not make use of. While our main results are remarkably stable and robust under the various specifications, our discussion below will use the estimates in columns 5 and 6, unless mentioned otherwise.¹⁶

Averaged across subjects, we see that teachers are 0.034σ /year more effective in teaching to their own gender relative to a student of the opposite gender compared to teachers of the other gender. In other words, female teachers are 0.034σ /year more effective in reducing the gender gap between girls and boys relative to male teachers. We find no negative effect on boys from being taught by female teachers relative to male teachers (β_3 is close to zero). We estimate that girls gain an extra 0.036σ /year when taught by female teachers instead of male teachers ($\beta_1 + \beta_3$) and that there would be no loss to the boys in the classroom. However, the net increase in annual test score gains from replacing a male teacher with a female one ($\lambda_g * \beta_1 + \beta_3$), which we estimate as 0.019σ /year, is not significant. Thus, while replacing a male teacher with a

female one on the margin is likely to benefit girls with no cost to boys, the magnitude of the positive effect on girls is small enough that the overall gain in test scores is not significant.¹⁷

Since female teachers systematically differ from their male counterparts in commonly observed characteristics used in teacher selection (Table 1 – Panel B), we examine the extent to which the shared gender effects estimated in Eq (1) can be attributed to female students being differentially affected by characteristics that are more commonly found in female teachers. Table 6 shows a series of regressions where we follow the specification in Eq (1), but include teacher characteristics and the interaction of this characteristic with student gender. These include teacher demographic characteristics that may be correlated with teaching effectiveness (such as education, training, contractual status, seniority, and salary), as well as teaching conditions (multi-grade teaching) and measures of teacher effort (absence). Doing so allows us to examine the extent to which we can attribute the mechanism for the positive β_1 found in Table 5 to observable teacher characteristics (that differ across male and female teachers) differentially affecting female students.

Panel A of Table 6 reports the key results with only the specified teacher characteristic and does not control for other teacher characteristics, while Panel B includes all the other teacher characteristics as controls. The estimates of β_1 are remarkably robust to including the student interactions with teacher characteristics that vary by teacher gender.¹⁸ In all cases, the estimate of the gain to a female student from switching to a female teacher ($\beta_1 + \beta_3$) is positive and significant (ranging from 0.027 to 0.04 σ /year), and we continue to find no negative effects on boys (we never reject $\beta_3 = 0$). When we include all teacher characteristics *and each of their interactions with student gender* (Panel B, Column 11), our estimate of β_1 falls slightly, but is not statistically different from previous estimates and continues to be statistically significant.

This suggests that the effects we find for female teachers closing the gender gap, though small, cannot be readily explained by other characteristics that could be used in teacher selection.

Thus, while we are unable to identify the specific mechanism behind the positive effect of female teacher on girls' learning gains, our results suggest that other characteristics used in teacher recruitment (and which systematically differ between male and female teachers) are not able to account for this effect. So if the goal of a policy maker is to reduce gender gaps in learning outcomes, our results suggest that teacher gender, even after holding observable characteristics constant, may be a useful summary statistic for other unobserved factors not used for recruiting teachers (such as empathy and role model effects) that contribute to better learning for girls.

B. Results by Subject

Table 7 breaks down the results by subject (Panels A and B) and also conducts formal tests of equality across subjects for the key parameters of interest (Panel C). Overall, we see that the patterns observed in Table 5 are seen consistently for both subjects, with no significant difference between them. There is a positive gender-match effect in both subjects ($\beta_1 > 0$) and the difference is not significant. Similarly, boys do no worse with female teachers and we cannot reject $\beta_3 = 0$ for either subject.

C. Robustness to an Alternative Specification

An alternative approach to our preferred specification is to identify the shared gender effect relative to each *teacher's* mean performance, which can be done by augmenting Eq (1) to include teacher fixed effects. This specification can be considered as an extension of Table 5 – Column 6, in that it controls for *all* time-invariant teacher characteristics, including those that we cannot

measure and include as controls (as we do in Table 5 – Column 6). We see in Table 8 that the estimates of β_1 from this specification are unchanged relative to those in Tables 5 through 7.

The unchanged estimates of β_1 when we include teacher fixed effects provides further confirmation of the stability of our core result regarding the significant role of female teachers in bridging test score genders gaps in primary school. This specification confirms that estimates of β_1 are not driven by unobserved characteristics correlated with female teachers that increase *overall* student achievement. However, we cannot estimate β_3 in this specification, which is essential to be able to estimate the policy impact of hiring more female teachers.¹⁹ Since a social planner would care about both increasing overall test scores as well as reducing the gender gap, we focus our discussions on our default specifications shown in Table 5, and present the results in Table 8 as a further robustness check on the matching result.²⁰

D. Attendance

We also study the impact of a teacher-student gender match on student attendance, using high-quality data on student attendance, measured using unannounced visits to schools (as opposed to using administrative data that is less reliable). We find no significant effect of a gender-match on student attendance (Table 9). We do find that female teachers are slightly more effective at increasing attendance overall (by approximately 0.6 percent), but there is no differential impact by student gender. This result is interesting because the rhetoric of hiring female teachers is often based on the belief that having female teachers increases the safety and comfort of girls in school and that their presence therefore encourages girls to attend school. Our results suggest, however, that the mechanism for the positive impact of a gender match on test scores is less likely to be due to effects on the extensive margin of increased school participation,

but more due to intensive margin increases in the effectiveness of classroom interactions between teachers and students.

This result could be reflecting the scenario where total primary school enrollment for both boys and girls is over 98 percent (Pratham 2012) and so the role of female teachers in increasing attendance of female students may be more limited in such a setting. Additionally, we observe attendance conditional upon enrollment rather than effects on enrollment into the school itself. Nevertheless, our results suggest that even after achieving gender parity in school enrollment, there may be continued benefits to a policy of preferred hiring of female teachers due to their effectiveness in reducing gender gaps in test scores.

E. Contribution of fewer female teachers in higher grades to growth in the gender gap

Finally, we calculate what proportion of the growing gender gap calculated in Table 2 can be attributed to girls being less likely to have a female teacher as they advance through primary school. Regressing the probability of a female teacher on the grade taught (with school fixed effects), we find that there is a four percentage point reduction in the probability of a student having a female teacher at each higher grade. Multiplying the reduced probability of a female teacher by the cost to girls of not having a female teacher in a given year ($\beta_1 + \beta_3$), and dividing this by the total annual increase in the test score gender gap (estimated in Table 2), we estimate that the reduced likelihood of female teachers in higher grades accounts for 9 percent of the annual growth in the gender gap in math and 21 percent in language. The fraction of the growing gender gap in language that is accounted for by this channel is higher than in math because the absolute magnitude of the annual growth in the gender gap is lower in language. Using estimates without school fixed effects, these figures would be 8 percent and 15 percent

respectively (because the overall trend in the gender gap is slightly larger without school fixed effects – see Table 2).

V. Conclusion

We study gender gaps in primary school learning outcomes in a low-income setting using one of the richest datasets on primary education in a developing country. We find that at the start of primary school, girls in rural public schools have a slight advantage in the local language (approximately 0.05σ) and are at par in math with the boys in the same schools. However, girls lose this advantage in both language (by $0.01\sigma/\text{year}$) and in math (by $0.02\sigma/\text{year}$) as they progress through the schooling system.

While these trends likely reflect a broad set of household, school, and social factors, one specific school-level policy that has been posited as a promising channel for mitigating gender gaps is the greater use of female teachers in low-income settings. Though this policy has been widely recommended and adopted, there has been very little well-identified evidence to support the claim. In this paper, we present some of the first well-identified empirical tests of this hypothesis in a developing country setting using an extremely rich data set that is representative of the rural public primary school system in the Indian state of Andhra Pradesh.

Our results suggest that female (and male) teachers are relatively more effective when teaching to their own gender, that learning for girls increases when they are taught by female teachers relative to male teachers, and that boys do not suffer adverse effects when taught by female teachers relative to male teachers, even when controlling for additional observable teacher characteristics.

Our results are similar to other studies which find positive effects on achievement via test scores in both high and low income countries. Our pattern on gender matching and the

magnitudes of our effects are very similar to those found in Dee (2007) in the US and by Rawal and Kingdon (2010) in India. Both these studies find positive effects for girls and no adverse effects for boys, with the shared female effect ranging from 0.03 - 0.06 standard deviations.

While we find suggestive evidence that the mechanism of impact is through more effective classroom interactions (as opposed to increased teacher-student contact time), our data does not allow us to explore the further granularity of the specific mechanisms through which shared gender may influence learning (such as role model effects, greater empathy, and closer identification between teachers and students of the same gender). We find no evidence that characteristics thought to be correlated with teaching effectiveness (education or training), service conditions (salary, multi-grade teaching), or teacher effort (absence) differentially affect girls and thus help explain our gender matching effects. Thus, even though we are unable to identify the mechanism for *why* female teachers differentially affect girls, we are at least able to show that the effect is not explained by differences (by teacher gender) in other common teacher recruitment characteristics. In other words, being female appears to be correlated with other unobservable characteristics that are correlated with classroom effectiveness, but which are not currently the basis for teacher recruitment.

From a policy perspective, our estimates suggest that expanding the hiring of female teachers - both at the margin of the current patterns of hiring (assuming that the marginal female teacher hired has the same characteristics as the average female teacher), and also when holding other typical recruitment characteristics constant, may be a useful tool for bridging gender gaps in learning levels and trajectories in primary schools, at no cost to boys. However, this result may not hold beyond primary school because the unobservable characteristics that are correlated with being female and teaching effectiveness may not be equally salient in higher grades.²¹

Further decomposing the reduced form effects of 'gender matching' and having a 'female' teacher could help in crafting more nuanced policies to more efficiently bridge gender gaps in schooling in developing countries.

1. Note that these estimates may be biased due to differential migration to private schools and differential absence on the day of the test by student gender. However, as we discuss later, these estimates are robust to bounding by reweighting to account for this concern.
2. Analogous to gender, studies in the United States have also looked at the effect of sharing the ethnicity of a teacher and have generally found positive effects on such educational outcomes as drop outs, pass rates, and grades at the community college level, and teacher perceptions and student achievement in school going children (Dee 2004; Dee 2005; Farlie, Hoffman, and Oreopoulos 2011). We find no similar effect on other important dimensions in the Indian context, particularly disadvantaged castes and minority religions. We do not focus on caste and religion because the fraction of teachers and students in the relevant categories are small (typically less than 20 percent) and as a result the fraction of 'matches' are usually less than 5 percent (and often much smaller), which makes the estimates less stable to the series of robustness checks that we use in this paper to ensure that the estimates of the 'match' are well-identified.
3. Thus, if this approach finds that a girl in eighth grade who has a female language teacher and a male math teacher does better in language, the interpretation of the point estimate is confounded by the possibility that the girl is also more likely to have had female language teachers in earlier grades (especially if teacher gender is correlated with subjects taught across grades, which is likely to be true).

4. The original state of AP was divided into two states on June 2, 2014. Since this division took place after our study, we use the term AP to refer to the original undivided state.

5. These interventions are described in Muralidharan and Sundararaman (2011).

6 The Appendix can be found online at <http://jhr.uwpress.org/>.

7. This is exactly analogous to thinking about the impact of attrition in randomized experiments, where treatment effects are typically not biased by attrition, as long as there is no differential attrition by treatment status (which in our case is the matching indicator).

8. Less than 3 percent of students with test scores have no recorded gender.

9. While there are a few observable differences between the boys and girls in the sample, including these in the estimation will only matter if there are differential interactions between these household characteristics and teacher gender across boys and girls. We verify that our results are robust to the inclusion of household characteristics (see Table A4), but prefer to not include household characteristics in our main estimating equations because doing so reduces the sample size by 30 percent and it is possible that the remaining sample may have some non-random attrition.

10. Note that this could also mean that girls with lower test scores are slightly more likely to stay in the sample, whereas their male counterparts are more likely to be absent from testing.

11. It is worth noting that the entire literature on gender gaps in test scores is based on samples of students who are tested in schools, and we know of no annual student-level panel data set on test scores (in any country) that can account for differential sorting into private schools and differential rates of dropping out or attendance. The only way to do this would to have a household panel on test scores using a representative sample of households – and no such data exists to the best of our knowledge. Thus, our reweighted estimates (that account for differential

attrition probability in our panel data) are likely to be the most reliable estimates of the evolution of gender gaps in test scores over grades in developing countries.

12. In the case of first grade where there is no lagged score (since there was no testing prior to enrolling in school), we set the normalized lagged score to zero. Our results on the impact of 'gender matching' on test score gains are unchanged if we drop first grade from the analysis.

13. Since the data are drawn from schools that were exposed to various experimentally-assigned programs, all estimates include dummy variables indicating the treatments assigned to the school. This turns out to not matter in practice because our main specifications of interest use school-fixed effects, which makes the treatment status of the school irrelevant for identification purposes.

14. As we note earlier, this is the approach used in most of the existing studies in this literature. However, an important weakness of this approach is that it is based on levels, and not value-added, which makes the estimates difficult to interpret without knowing the gender of teachers for each subject in earlier grades.

15. We avoid using a student fixed effects estimate because the identifying variation in a specification with student fixed effects would come from changes in teacher gender in different grades. However, girls having higher value-added in lower grades and female teachers being more likely to be assigned to lower grades would create an upward bias in the 'matching estimate'.

16. All coefficients and tests continue to be of similar statistical significance under specifications with standard errors clustered at the school or teacher level.

17. In other words, since the fraction of girls in the sample is roughly half ($\lambda_g = 0.51$), the positive effect of female teachers on girls is not large enough for ‘half’ of this effect to be significant.

18. In the interest of space, we only show these results for the subset of characteristics that are significantly different across teacher gender (see Table 1 – Panel B). The estimate of β_1 is unchanged and significant for interactions with all other teacher characteristics in Table 1 – Panel B (such as religion and caste) as well.

19. A positive and significant β_1 is possible even if female teachers reduced gender gaps by being less effective at teaching boys, while being no more effective at teaching girls than male teachers (or even less so). Thus it is essential to use β_3 in conjunction with β_1 to estimate the overall impact of hiring more female teachers.

20. Note that identification concerns regarding β_3 are also addressed by the results in Table 3, where we see no significant difference in the initial test scores of students assigned to a female teacher in any of the six columns. Also, our default specification uses school-grade fixed effects, mitigating concerns of omitted variables correlated with teacher gender and student test-score gains both across and within schools across grades. Table 8 provides further suggestive evidence that the estimates of β_1 , β_2 , and β_3 in Tables 5 and 6 are unbiased, because β_1 , and β_2 are unchanged when we include teacher fixed effects.

21. We do not find higher effects of female teachers on female students in earlier grades, suggesting that at least within primary school, there is no evidence of a declining ‘gender match’ or ‘female teacher’ effect from the first to the fifth grade. Nevertheless, we cannot extrapolate these results beyond primary school.

References

- Antecol, Heather, Ozkan Eren, and Serkan Ozbeklik. 2015. "The Effect of Teacher Gender on Student Achievement in Primary School." *Journal of Labor Economics*. Forthcoming.
- Bettinger, Eric P., and Bridget Terry Long. 2005. "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students." *American Economic Review* 95(2): 152-157.
- Bharadwaj, Prashant, Katrina Loken, and Christopher Neilson. 2013. "Early Life Health Interventions and Academic Achievement." *American Economic Review* 103(5): 1862 – 1891.
- Bharadwaj, Prashant, Giacomo De Giorgi, David Hansen, and Christopher Neilson. 2012. "The Gender Gap in Mathematics: Evidence from Low and Middle Income Countries." NBER Working Paper.
- Carrell, Scott E., Marianne E. Page, James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *Quarterly Journal of Economics* 125(3): 1101 – 1144.
- Carrington, Bruce, Peter Tymms, and Christine Merrell. 2008. "Role Models, School Improvement and the 'Gender Gap' – Do Men Bring Out the Best in Boys and Women Bring Out the Best in Girls?" *British Educational Research Journal* 34(3): 315 – 327.
- Dee, Thomas. 2004. "Teachers, Race and Student Achievement in a Randomized Experiment." *The Review of Economics and Statistics* 86(1): 195 – 210.
- _____. 2005. "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?" *American Economic Review* 95(2): 158 – 165.

- _____. 2007. "Teachers and the Gender Gaps in Student Achievement." *Journal of Human Resources* 42(3): 528 – 554.
- Driessen, Geert. 2007. "The Feminization of Primary Education: Effects of Teachers' Sex on Pupil Achievement, Attitudes and Behavior." *Review of Education* 53: 183 – 203.
- Ehrenberg, Ronald G., Dan D. Goldhaber, and Dominic J. Brewer. 1995. "Do Teachers' Race, Gender, and Ethnicity Matter? Evidence from the National Educational Longitudinal Study of 1988." *Industrial and Labor Relations Review* 48(3): 547 – 561.
- Fairlie, Robert W., Florian Hoffmann, and Philip Oreopoulos. 2014. "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom." *American Economic Review* 104(8): 2567 – 91.
- Fryer, Roland, and Steven D. Levitt. 2010. "An Empirical Analysis of the Gender Gap in Mathematics." *American Economic Journal: Applied Economics* 2(2): 210 – 40.
- Government of India. 2008. *Eleventh Five Year Plan 2007 – 2012*. New Delhi: Oxford University Press.
- Hausmann, Ricardo, Laura D. Tyson, and Saadia Zahidi. 2012. *The Global Gender Gap Report 2012*. Geneva: World Economic Forum.
- Herz, Barbara, Gene B. Sperling. 2004. *What Works in Girls' Education*. Council on Foreign Relations.
- Hoffman, Florian, Philip Oreopoulos. 2009. "A Professor Like Me: The Influence of Instructor Gender on College Achievement." *Journal of Human Resources* 44(2): 479 – 494.
- Holmlund, Helena, Krister Sund. 2008. "Is the Gender Gap in School Performance Affected by the Sex of the Teacher?" *Labour Economics* 15(1): 37 – 53.

- Lahelma, Elina. 2000. "Lack of Male Teachers: A Problem for Students or Teachers?" *Pedagogy, Culture and Society* 8(2): 173 – 86.
- Marsh, Herbert W., Andrew J. Martin, and Jacqueline H.S. Chend. 2008. "A Multilevel Perspective on Gender in Classroom Motivation and Climate: Potential Benefits of Male Teachers for Boys?" *Journal of Educational Psychology* 100(1): 78 – 95.
- Muralidharan, Karthik, and Nishith Prakash. 2013. "Cycling to School: Increasing Secondary School Enrollment for Girls in India." NBER Working Paper.
- Muralidharan, Karthik, Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119(1): 39 – 77.
- Neugebauer, Martin, Marcel Helbig, and Andreas Landmann. 2011. "Unmasking the Myth of the Same-Sex Teacher Advantage." *European Sociological Review* 27(5): 669 – 689.
- Nixon, Lucia, and Michael Robinson. 1999. "The Educational Attainment of Young Women: Role Model Effects of Female High School Faculty." *Demography* 36(2): 185 – 194.
- OECD. 2010. *PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume 1)*. OECD.
- Ouazad, Amine, and Lionel Page. 2012. *Students' Perceptions of Teachers Biases: Experimental Economics in Schools*. London: London School of Economics.
- Pratham. 2012. *Annual Status of Education Report*. Pratham.
- Rawal, Shenila, and Geeta Kingdon. 2010. "Akin to My Teacher: Does Caste, Religious, or Gender Distance Between Student and Teacher Matter? Some Evidence from India." University of College London Working Paper.
- Rehman, Naseem-Ur Rehman. 2008. "Yemen Makes Progress in Girls' Education with UNICEF – Supported Literacy Programmes." *UNICEF Newslines*. March 31.

Slavin, Patrick. 2006. "Liberia Launches Girls' Education National Policy with Support from UNICEF." *UNICEF Newslines*. April 18.

Stacki, Sandra. 2002. *Women Teachers Empowered in India: Teacher Training Through a Gender Lens*. New York: UNICEF.

UNDG. 2010. *Thematic Paper on MDG3: Promote Gender Equality and Empower Women*. United Nations Development Group.

UNESCO. 2010. Institute for Statistics. *UNESCO*. 31 August 2013.

_____. 2012. *UNESCO eAtlas of Gender Equality in Education*. Paris: UNESCO.

Winters, Marcus A., Robert C. Haight, Thomas T. Swaim, and Kararzyna A. Pickering. 2013.

"The Effect of Same-Gender Teacher Assignment on Student-Achievement in the Elementary and Secondary Grades: Evidence from Panel Data." *Economics of Education Review* 34: 69 – 7.

Table 1
Summary Statistics by Gender

	Panel A: Students				
	Obs	Mean	Male	Female	Female - Male
Female	94,599	0.51			
Literate Father	66,511	0.59	0.58	0.60	0.0185***
Literate Mother	66,827	0.44	0.43	0.45	0.0199***
Proper House	66,851	0.31	0.31	0.32	0.0098***
Has Toilet	66,974	0.29	0.28	0.29	0.0106***
	Panel B: Teachers				
	Obs	Mean	Male	Female	Female - Male
Female	2,680	0.46			
Head Teacher	2,680	0.29	0.38	0.18	-0.195***
Regular Teacher	2,680	0.50	0.50	0.51	0.0141
Contract Teacher	2,680	0.19	0.12	0.27	0.157***
Completed Education: 12th Pass	2,680	0.93	0.96	0.89	-0.0696***
Completed Education: Masters	2,680	0.23	0.27	0.17	-0.0964***
Has Teacher Training	2,661	0.83	0.91	0.74	-0.166***
Native to Village	2,679	0.23	0.18	0.30	0.128***
Married	2,676	0.81	0.85	0.77	-0.0762***
Active in Union	2,674	0.18	0.28	0.07	-0.202***
Salary (monthly)	2,674	9,560.21	10,697.00	8,209.46	-2487.5***
Age	2,660	36.91	39.54	33.75	-5.791***
Years Experience	2,285	12.95	14.47	11.08	-3.389***

Teacher Absence	2,666	0.19	0.20	0.18	-0.0135**
Multi-grade Classroom	2,680	0.46	0.48	0.44	-0.0386**
Classroom Enrollment	2,680	23.22	22.87	23.65	0.778

Notes: (1) All variables are binary indicators, except for salary which ranges from 300 to 38400 (with a standard deviation of 5776), age which ranges from 12 to 58 (with a standard deviation of 9.76), and years of experience which ranges from one to 42 (with a standard deviation of 7.94). (2) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2
Learning Gaps by Gender and
Grade

	Dependent Variables: Normalized Test Score (Within Grade)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled Across Subjects		Math		Telugu	
Panel A: Gender Differentials in Test Scores by Grade						
Female (Grade 1)	0.0279** (0.0122)	0.0206** (0.0101)	0.00238 (0.0135)	-0.00376 (0.0110)	0.0531*** (0.0127)	0.0448*** (0.0109)
Observations	66,660	66,660	33,187	33,187	33,473	33,473
Female (Grade 2)	0.00526 (0.0114)	0.00571 (0.00929)	-0.0271** (0.0117)	-0.0242** (0.00985)	0.0376*** (0.0122)	0.0354*** (0.00993)
Observations	70,953	70,953	35,453	35,453	35,500	35,500
Female (Grade 3)	-0.0217* (0.0118)	-0.0225** (0.00932)	-0.0569*** (0.0120)	-0.057*** (0.00984)	0.0136 (0.0128)	0.0121 (0.0103)
Observations	74,715	74,715	37,349	37,349	37,366	37,366
Female (Grade 4)	-0.044*** (0.0120)	-0.038*** (0.00947)	-0.0956*** (0.0122)	-0.088*** (0.0101)	0.00709 (0.0130)	0.0122 (0.0104)
Observations	79,972	79,972	39,973	39,973	39,999	39,999
Female (Grade 5)	-0.0262** (0.0115)	-0.0209** (0.00900)	-0.0749*** (0.0123)	-0.067*** (0.00955)	0.0225* (0.0123)	0.0254** (0.0100)
Observations	85,572	85,572	42,777	42,777	42,795	42,795
Panel B: Grade Trends in Gender Differentials in Test Scores Unweighted						
Female	0.0407***	0.0365***	0.0224	0.0192	0.0586***	0.0534***

	(0.0142)	(0.0116)	(0.0153)	(0.0125)	(0.0148)	(0.0126)
Female*Grade	-0.0170*** (0.00410)	-0.0154*** (0.00335)	-0.024*** (0.00443)	-0.022*** (0.00356)	-0.0102** (0.00435)	-0.0084** (0.00369)
Observations	304,410	304,410	151,785	151,785	152,625	152,625
School Fixed Effects	No	Yes	No	Yes	No	Yes
Panel C: Grade Trends in Gender Differentials in Test Scores Weighted						
Female	0.0447*** (0.0142)	0.0397*** (0.0116)	0.0273* (0.0152)	0.0229* (0.0124)	0.0617*** (0.0148)	0.0560*** (0.0126)
Female*Grade	-0.0165*** (0.00407)	-0.0147*** (0.00335)	-0.023*** (0.00440)	-0.022*** (0.00355)	-0.00914** (0.00431)	-0.00742** (0.00369)
Observations	304,410	304,410	151,785	151,785	152,625	152,625
School Fixed Effects	No	Yes	No	Yes	No	Yes

Notes: (1) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including school fixed effects, and are clustered at the school-year level for OLS regressions including school fixed effects. (2) Panel B and C includes only observations from grade one and those which have a recorded test score from the previous year. (3) Panel C weights the regression by a predicted probability of remaining in the sample as a function of the previous year's test score and gender. (4) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3
 Characteristics of Classrooms Assigned to Female Teachers

	Dependent Variable: Classroom Has a Female Teacher					
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion of Female Students	-0.00067 (0.0338)	-0.0121 (0.0323)	0.0011 (0.0215)	-0.0103 (0.0202)	0.0052 (0.0205)	-0.0091 (0.0197)
Grade 1	0.0156 (0.0163)	0.064*** (0.0153)	0.024** (0.0123)	0.066*** (0.0117)		
Grade 2	0.0228 (0.0150)	0.049*** (0.0141)	0.028** (0.0120)	0.046*** (0.0113)		
Grade 4	-0.067*** (0.0150)	-0.034*** (0.0146)	-0.068*** (0.0122)	-0.036*** (0.0117)		
Grade 5	-0.140*** (0.0170)	-0.063*** (0.0163)	-0.134*** (0.0131)	-0.054*** (0.0126)		
Test Score of Incoming Cohort of Male Students	-0.0142 (0.0228)	-0.0111 (0.0214)	-0.0015 (0.0136)	-0.00047 (0.0131)	-0.0091 (0.0141)	-0.0038 (0.0139)
Test Score of Incoming Cohort of Female Students	0.0189 (0.0191)	0.0068 (0.0188)	-0.0070 (0.0123)	-0.0048 (0.0118)	0.0017 (0.0130)	0.0057 (0.0127)
Observations	10,974	9,641	10,974	9,641	10,974	9,641
Teacher Characteristics	No	Yes	No	Yes	No	Yes
School Fixed Effects	No	No	Yes	Yes	No	No
School*Grade Fixed Effects	No	No	No	No	Yes	Yes
Boys' Test Score = Girls' Test Score	0.3168	0.5708	0.7932	0.8298	0.6117	0.6483

(p-value)

Notes: (1) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multi-grade class. (2) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the school-year level for OLS regressions including fixed effects. (3) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4
Gender Differentials in Learning Trajectories from Lower to Higher Grades

	Dependent Variable: Normalized Test Scores					
	(1)	(2)	(3)	(4)	(5)	(6)
	Pooled Across Subject		Math		Language	
Female	0.0251** (0.0120)	0.0255** (0.0112)	0.00464 (0.0132)	0.00667 (0.0121)	0.0464*** (0.0125)	0.0453*** (0.0122)
Female*Grade	-0.00624* (0.00322)	-0.00725** (0.00302)	-0.00830** (0.00368)	-0.0106*** (0.00334)	-0.00563* (0.00339)	-0.00562* (0.00332)
Observations	304,410	304,410	151,785	151,785	152,625	152,625
School Fixed Effects	No	Yes	No	Yes	No	Yes

Notes: (1) Regressions include student's previous year's test score as an independent variable. (2) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including school fixed effects, and are clustered at the school-year level for OLS regressions including school fixed effects. (3) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5

Impact of Female Teachers on the Learning Gains of Female Students (Pooled Across Math and Language)

	Dependent Variable: Normalized Test Scores					
	(1)	(2)	(3)	(4)	(5)	(6)
(β_1) Female Student * Female Teacher	0.0383*** (0.00997)	0.0362*** (0.00886)	0.0354*** (0.00827)	0.0350*** (0.00887)	0.0343*** (0.00826)	0.0347*** (0.00868)
(β_2) Female Student	-0.0120* (0.00676)	-0.0140** (0.00602)	-0.0126** (0.00563)			
(β_3) Female Teacher	-0.0154 (0.0188)	-0.00344 (0.0148)	0.000700 (0.0155)	0.00212 (0.0149)	0.00132 (0.0155)	-0.00305 (0.0180)
$\beta_1 + \beta_3$	0.023	0.033**	0.036**	0.037**	0.036**	0.032***
F-statistic ($H_0: \beta_1 + \beta_3 = 0$)	1.575	5.222	5.683	6.448	5.500	3.157
$\lambda_g * \beta_1 + \beta_3$	0.004	0.015	0.019	0.020	0.019	0.015
F-statistic ($H_0: \lambda_g * \beta_1 + \beta_3 = 0$)	0.054	1.169	1.613	2.012	1.618	0.708
Observations	268,548	268,548	268,548	268,548	268,548	235,022
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes
Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes

Notes: (1) Regressions include student's previous year's test score as an independent variable. (2) "Teacher Characteristics"

are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education,

training, native to school location, marital status, union status, and a multi-grade class. (3) Standard errors (in parentheses)

are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the school-year level

for OLS regressions including fixed effects. (4) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6

Heterogeneous Effects on Test Score Gains of Girls by Teacher Characteristics and Teacher Gender

Teacher Characteristic:	Dependent Variable: Normalized Test Scores										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Head Teacher	Contract Teacher	Completed 12th	Teacher Training	Native to Village	Active in Union	Salary (per Rs.10,000)	Experience	Absence	Teaches Multi-grade Classroom	All Teacher Characteristics
Panel A: Excludes Additional Teacher Correlates											
(β_1) Female Student * Female Teacher	0.031*** (0.00852)	0.034*** (0.00836)	0.035*** (0.00827)	0.035*** (0.00836)	0.034*** (0.00829)	0.032*** (0.0086)	0.036*** (0.00848)	0.030*** (0.00888)	0.035*** (0.0083)	0.034*** (0.00827)	
(β_3) Female Teacher	-0.00228 (0.0157)	0.00243 (0.0157)	-0.000149 (0.0154)	0.000203 (0.0157)	-0.00156 (0.0155)	0.00834 (0.0158)	-0.00201 (0.0158)	-0.000133 (0.0174)	-0.00350 (0.0155)	0.000449 (0.0154)	
(δ_1) Female Student *Characteristic	-0.0179* (0.00968)	0.000603 (0.0137)	0.0260 (0.0198)	-0.00151 (0.0141)	0.00786 (0.0110)	-0.0132 (0.0112)	0.0077 (0.00839)	-0.000939 (0.00060)	-0.00428 (0.0200)	-0.030*** (0.00878)	
(δ_3) Teacher Characteristic	-0.0185 (0.0161)	-0.00902 (0.0238)	-0.0423 (0.0360)	0.0102 (0.0247)	0.00785 (0.0181)	0.0413** (0.0178)	-0.0177 (0.0153)	-0.00165 (0.00113)	-0.067** (0.0304)	-0.0162 (0.0182)	
$\beta_1 + \beta_3$ F-stat ($H_0: \beta_1 + \beta_3 = 0$)	0.029* 3.47	0.037** 5.70	0.035** 5.35	0.035** 5.14	0.032** 4.53	0.040*** 6.68	0.034** 4.83	0.030* 3.10	0.032** 4.33	0.034** 5.09	
$\lambda_g * \beta_1 + \beta_3$ F-stat ($H_0: \lambda_g * \beta_1 + \beta_3 = 0$)	0.014 0.82	0.020 1.76	0.018 1.45	0.018 1.42	0.016 1.13	0.025 2.66	0.016 1.17	0.015 0.84	0.014 0.95	0.017 1.43	
Observations	268,548	268,548	268,548	267,475	268,482	268,318	268,215	239,932	264,581	268,264	
Panel B: Includes Additional Teacher Correlates											
(β_1) Female Student * Female Teacher	0.032*** (0.0089)	0.034*** (0.0088)	0.036*** (0.0087)	0.034*** (0.0087)	0.034*** (0.0087)	0.032*** (0.0091)	0.035*** (0.0089)	0.032*** (0.0090)	0.035*** (0.0087)	0.033*** (0.00868)	0.026*** (0.0095)
(β_3) Female Teacher	-0.00148 (0.0180)	-0.00241 (0.0180)	-0.00361 (0.0180)	-0.00254 (0.0180)	-0.00282 (0.0180)	-0.00156 (0.0179)	-0.00342 (0.0180)	-0.00174 (0.0180)	-0.00305 (0.0180)	-0.00249 (0.0180)	0.00152 (0.0181)
(δ_1) Female Student *Characteristic	0.0972 (0.0983)	0.0179 (0.0815)	-0.0449 (0.0534)	0.0455 (0.0487)	0.0259 (0.0251)	0.0430** (0.0188)	-0.00291 (0.0262)	0.000493 (0.00179)	-0.071** (0.0334)	-0.0172 (0.0196)	

(δ_3) Teacher Characteristic	-0.0197*	0.0169	0.0372	-0.0146	0.0118	-0.0141	0.00453	-0.000854	-0.00021	-0.032***	
	(0.0101)	(0.0160)	(0.0235)	(0.0169)	(0.0123)	(0.0117)	-0.00914	(0.000613)	(0.0213)	(0.00925)	
$\beta_1 + \beta_3$	0.030*	0.031*	0.032*	0.031*	0.031*	0.030*	0.032*	0.030*	0.032*	0.031*	0.027
F-stat ($H_0: \beta_1 + \beta_3 = 0$)	2.86	3.06	3.24	3.09	3.12	2.87	3.21	2.93	3.16	3.02	2.34
$\lambda_g * \beta_1 + \beta_3$	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015	0.015
F-stat ($H_0: \lambda_g * \beta_1 + \beta_3 = 0$)	0.71	0.71	0.70	0.71	0.71	0.72	0.71	0.72	0.71	0.70	0.71
Observations	235,022	235,022	235,022	235,022	235,022	235,022	235,022	235,022	235,022	235,022	235,022

Notes: (1) Regressors include student's lagged normalized test score, school*grade fixed effects and grade fixed effects by student gender (Specification from Column 5 and 6 of Table

5). (2) "Teacher Characteristics" are salary per 10,000 Rs., age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and an indicator for teaching a multi-grade class. (3) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the school-year level for OLS regressions including fixed effects. (4) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7
Impact of Female Teachers on the Learning Gains of Female Students by Subject

	Dependent Variable: Normalized Test Scores					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Math					
(β_1) Female Student * Female Teacher	0.0338*** (0.0106)	0.0335*** (0.00992)	0.0336*** (0.00940)	0.0312*** (0.00994)	0.0315*** (0.00940)	0.0302*** (0.00988)
(β_2) Female Student	-0.0374*** (0.00736)	-0.0428*** (0.00669)	-0.0408*** (0.00633)			
(β_3) Female Teacher	-0.0139 (0.0173)	0.00240 (0.0164)	0.00806 (0.0175)	0.00916 (0.0166)	0.00917 (0.0175)	0.00924 (0.0205)
$\beta_1 + \beta_3$	0.020	0.036**	0.042**	0.040**	0.041**	0.039*
F-stat ($H_0: \beta_1 + \beta_3 = 0$)	1.365	5.000	5.760	6.047	5.494	3.714
$\lambda_g * \beta_1 + \beta_3$	0.003	0.019	0.025	0.025	0.025	0.025
F-stat ($H_0: \lambda_g * \beta_1 + \beta_3 = 0$)	0.040	1.583	2.245	2.538	2.257	1.531
Observations	133,907	133,907	133,907	133,907	133,907	117,205
	Panel B: Language (Telugu)					
(β_1) Female Student * Female Teacher	0.0429*** (0.0104)	0.0393*** (0.00979)	0.0373*** (0.00915)	0.0385*** (0.00979)	0.0364*** (0.00917)	0.0392*** (0.00971)
(β_2) Female Student	0.00971 (0.00718)	0.0104 (0.00669)	0.0113* (0.00629)			
(β_3) Female Teacher	-0.0174 (0.0154)	-0.00858 (0.0154)	-0.00531 (0.0159)	-0.00389 (0.0155)	-0.00485 (0.0159)	-0.0140 (0.0184)
$\beta_1 + \beta_3$	0.026*	0.031**	0.032**	0.035**	0.032**	0.025
F-stat ($H_0: \beta_1 + \beta_3 = 0$)	2.873	4.128	4.204	5.097	4.091	1.922

$\lambda_g * \beta_1 + \beta_3$	0.005	0.011	0.014	0.016	0.014	0.006
F-stat ($H_0: \lambda_g * \beta_1 + \beta_3 = 0$)	0.098	0.623	0.822	1.156	0.825	0.115
Observations	134,641	134,641	134,641	134,641	134,641	117,817
Panel C: Subject Differences (Math - Language)						
$\beta_{1M} - \beta_{1L}$	-0.009	-0.006	-0.004	-0.007	-0.005	-0.009
χ^2 [$H_0: \beta_{1M} - \beta_{1L} = 0$]	1.30	0.45	0.20	0.72	0.35	1.04
$\beta_{3M} - \beta_{3L}$	0.004	0.011	0.013	0.013	0.014	0.023
χ^2 [$H_0: \beta_{3M} - \beta_{3L} = 0$]	0.11	0.88	1.20	1.20	1.31	2.59
$(\beta_{1M} + \beta_{3M}) - (\beta_{1L} + \beta_{3L})$	-0.006	0.005	0.010	0.006	0.009	0.014
χ^2 [$H_0: (\beta_{1M} + \beta_{3M}) - (\beta_{1L} + \beta_{3L}) = 0$]	0.60	0.19	0.59	0.22	0.53	0.95
$(\lambda_g * \beta_{1M} + \beta_{3M}) - (\lambda_g * \beta_{1L} + \beta_{3L})$	-0.001	0.008	0.011	0.009	0.012	0.019
χ^2 [$H_0: (\lambda_g * \beta_{1M} + \beta_{3M}) - (\lambda_g * \beta_{1L} + \beta_{3L}) = 0$]	0.02	0.53	0.96	0.68	0.97	1.82
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes
Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes

Notes: (1) Regressions include student's previous year's test score as an independent variable. (2) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multi-grade class. (3) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the school-year level for OLS regressions including fixed effects. (4) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8

Impact of Female Teachers on the Learning Gains of Female Students Using Teacher Fixed Effects

	Dependent Variable: Normalized Test Scores					
	(1)	(2)	(3)	(4)	(5)	(6)
	Across Subjects		Math		Language	
(β_1) Female Student * Female Teacher	0.0357*** (0.00807)	0.0347*** (0.00807)	0.0337*** (0.00919)	0.0314*** (0.00921)	0.0377*** (0.00909)	0.0375*** (0.00909)
(β_2) Female Student	-0.0127** (0.00554)		-0.0420*** (0.00628)		0.0122* (0.00624)	
Observations	268,548	268,548	133,907	133,907	134,641	134,641
Teacher Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Grade Fixed Effects by Student Gender	No	Yes	No	Yes	No	Yes

Notes: (1) Regressions include student's previous year's test score as an independent variable. (2) Standard errors (in parentheses)

are clustered at the school-year level. (3) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9
Impact of Female Teachers on the Attendance of Female Students

	Dependent Variable: Student Attendance					
	(1)	(2)	(3)	(4)	(5)	(6)
(β_1) Female Student * Female Teacher	-0.00266 (0.00349)	-0.00380 (0.00321)	-0.00346 (0.00318)	-0.00326 (0.00322)	-0.00288 (0.00321)	-0.00178 (0.00345)
(β_2) Female Student	0.00759*** (0.00264)	0.00740*** (0.00219)	0.00679*** (0.00216)			
(β_3) Female Teacher	0.000113 (0.00461)	-0.00372 (0.00342)	0.00752* (0.00385)	0.00375 (0.00344)	0.00720* (0.00386)	0.00588 (0.00437)
Observations	148,791	148,791	148,791	148,791	148,791	129,890
Male Student with Male Teacher Mean	0.777	0.777	0.777	0.777	0.777	0.777
$\beta_1 + \beta_3$	-0.003	-0.008**	0.004	0.000	0.004	0.004
F-stat ($H_0: \beta_1 + \beta_3 = 0$)	0.346	5.063	1.175	0.021	1.321	0.936
$\lambda_g * \beta_1 + \beta_3$	-0.002	-0.011*	0.012*	0.004	0.012*	0.010
F-stat ($H_0: \lambda_g * \beta_1 + \beta_3 = 0$)	0.087	3.558	2.815	0.507	2.782	1.601
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes
Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes

Notes: (1) Attendance is calculated as the average of the indicator of whether the student was present or not on the day of 2 to 6 visits per year. (2) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multi-grade class. (3) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the school-year level for OLS regressions including fixed effects. Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix 1

Data and Attrition

The data used in this paper were collected over five school years from 2005-06 to 2009-10 from a representative sample of government-run rural primary schools in the Indian state of Andhra Pradesh (AP). Since primary school consists of grades one through five, a total of nine cohorts of students are present in our data (with the oldest cohort being in grade five in year one of the project, and the youngest cohort being in grade one in year five). Table A1 shows the number of student observations by grade and year in our core estimation sample.

For students in grades two through five, the estimating sample includes only those who have a test score in the current grade/year and have a test score from the previous grade/year (which is needed to estimate value-added). For first grade students, we include all those who have a test score in grade one and set the normalized lagged test score to zero since there is no previous test (the estimates in Tables 5 through 8 are unchanged even if we exclude first grade). For grades two through five, field teams conducted two rounds of testing at the end of each year (the first test covered competencies from the previous year and the second test covered current year competencies). Since student attendance rates are approximately 70 percent, having two rounds of testing helps considerably with reducing attrition from the sample.¹ However, there is only one round of testing at the end of first grade (since there are no previous grade competencies to be covered). Thus, the grade two sample in any year is smaller than the other grades.²

This sample is further limited to observations for which we have student gender data (97 percent) and for specifications that include the teacher characteristics, the sample is restricted to cases where teacher interviews were conducted (which is 88 percent of the sample conditional on having student test data and student and teacher gender data for the year).

Moving across a row in Table A1 (over years), we observe a reduction in student observations. This is because the share of private school enrollment is growing considerably in rural Andhra Pradesh (Pratham 2012) and fewer students are entering the public school system over time. Table A2 tests whether entering cohorts over time differ in relative ability by student gender. We find no differences in grade one test scores over time by student gender suggesting that the ability of girls relative to boys is not changing over time for the later entering cohorts. Thus, our estimates of the gender gap or of the impact of students sharing a teacher's gender are unlikely to be affected by the changing cohort sizes and composition over time.

We next review how attrition from the sample will affect our estimates and interpretation of the gender gap and the effect of 'gender matching'. Attrition is defined as the fraction of students in a given year who are in the potential estimation sample (which comprises of all students who have a valid test score for the previous year), but are not in the final sample because they were absent from the end of year test and thus have no recorded test score for the current year. Grade one students are not included in the attrition analysis because they do not have a test score from the previous year, and we therefore cannot define attrition for first grade. As mentioned earlier, all the results in Tables 5 through 8 are robust to excluding grade one.

From our analysis on student attendance (Table 9), we know that girls are less likely to be absent from school on any given school day. Similarly, we find that girls have lower attrition (3 percent) in the sample used for the value-added calculations (Table A3). But we also see that there is no effect of a student having the same gender as their teacher on the probability of attrition. Thus, our main estimates (presented in Tables 5 through 8) are unlikely to be biased due to the lower attrition of girls from our estimation sample. Furthermore, the differential attrition by student gender will only change our interpretation of the gender matching effect if the

students who attrite are *differentially affected* by shared teacher gender, which is unlikely given the lack of any effect of gender matching on either student attendance (Table 9 – Columns 5 and 6), or on the probability of taking an end of year test conditional on having taken the test at the end of the previous school year (Table A3 – Columns 5 and 6).

1. Student scores are first normalized with respect to each test and then averaged across the two tests to provide a valid normalized test score for any student who took at least one of the two tests.
2. First grade has the highest number of missing students in the end-line, but does not require a baseline; and grades three through five have the benefit of fewer missing data points since they are less likely to have missing test score data from the previous year (where there would have been two rounds of testing).

Table A1
Estimating Sample by Year and Grade

	Year 1	Year 2	Year 3	Year 4	Year 5
Grade 1	14,011	13,030	11,332	11,150	9,194
Grade 2	10,286	8,021	8,322	6,778	6,162
Grade 3	11,496	10,381	10,372	9,757	8,276
Grade 4	14,119	11,430	10,702	11,010	9,711
Grade 5	15,415	14,024	11,801	11,295	10,473

Table A2
Entering Cohorts by Gender

	Dependent Variable: Normalized Test Score	
	(1)	(2)
Female Student	-0.00135 (0.0261)	-0.0000674 (0.0229)
Year	-0.00532 (0.0120)	-0.00540 (0.00978)
Female Student * Year	0.0102 (0.00832)	0.00725 (0.00726)
Observations	66,660	66,660
School Fixed Effects	No	Yes

Notes: (1) Sample limited to students in Grade 1. (2) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (3) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3
Attrition and Gender

	Dependent Variable: Indicator of Attrition					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Attrition by Gender Matching						
Female Student * Female Teacher	-0.00657 (0.00527)	-0.00620 (0.00488)	-0.00871* (0.00474)	-0.00322 (0.00480)	-0.00546 (0.00475)	-0.00546 (0.00507)
Female Student	-0.0322*** (0.00380)	-0.0335*** (0.00334)	-0.0308*** (0.00323)			
Female Teacher	0.00870 (0.00687)	0.0155** (0.00604)	0.00476 (0.00554)	-0.00555 (0.00530)	0.00304 (0.00555)	0.00468 (0.00625)
Observations	132,671	132,671	132,671	132,671	132,671	116,473
Male Student Attrition Mean	0.223	0.223	0.223	0.223	0.223	0.223
Female Student Attrition Mean	0.188	0.188	0.188	0.188	0.188	0.188
Panel B: Attrition by Ability and Student Gender						
Lag Test Score	-0.0365*** (0.00288)	-0.0377*** (0.00230)	-0.0351*** (0.00212)	-0.0363*** (0.00215)	-0.0350*** (0.00212)	-0.0354*** (0.00234)
Female Student	-0.0363*** (0.00281)	-0.0377*** (0.00234)	-0.0361*** (0.00226)	-0.0172*** (0.00393)	-0.0177*** (0.00388)	-0.0223*** (0.00437)
Female Student * Lag Test Score	0.00837*** (0.00277)	0.00827*** (0.00243)	0.00476** (0.00233)	0.00567** (0.00237)	0.00481** (0.00232)	0.00638** (0.00264)
Observations	149,970	149,970	149,970	149,970	149,970	115,592
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes

Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes
---------------------------------------	----	----	----	-----	-----	-----

Notes: (1) Student Attrition is calculated as an indicator for being absent for the test in a given year and having taken the test the preceding year. (2) Grade 1 students are excluded because they do not have a test score prior to enrollment in school. (3) Year 1 students who drop out of the sample in the first year are excluded. (4) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (5) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4
 Impact of Female Teachers on the Learning Gains of Female Students (Pooled Across Math and Language)
 with Student Characteristics (Table 5 with Student Characteristics)

	Dependent Variable: Normalized Test Scores					
	(1)	(2)	(3)	(4)	(5)	(6)
(β_1) Female Student * Female Teacher	0.0490*** (0.0108)	0.0463*** (0.00998)	0.0439*** (0.00935)	0.0450*** (0.00995)	0.0425*** (0.00929)	0.0433*** (0.00976)
(β_2) Female Student	-0.0297*** (0.00756)	-0.0282*** (0.00669)	-0.0240*** (0.00622)			
(β_3) Female Teacher	-0.0260 (0.0189)	-0.0127 (0.0155)	-0.0119 (0.0161)	-0.0112 (0.0155)	-0.0112 (0.0160)	-0.0185 (0.0187)
$\beta_1 + \beta_3$	0.023	0.034**	0.032**	0.034**	0.031**	0.025
F-stat ($H_0: \beta_1 + \beta_3 = 0$)	1.71	5.18	4.28	5.11	4.10	1.83
$\lambda_g * \beta_1 + \beta_3$	-0.001	0.011	0.010	0.012	0.010	0.004
F-stat ($H_0: \lambda_g * \beta_1 + \beta_3 = 0$)	0.00	0.58	0.48	0.66	0.48	0.04
Observations	207,819	207,819	207,819	207,819	207,819	181,468
Student Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes
Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes

Notes: (1) Regressions include student's previous year's test score as an independent variable and all student characteristics listed in Table 1, Panel A. (2) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multi-grade class. (3) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the school-year level for OLS regressions including fixed effects. (4) Significance levels are as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.