The impact of student-teacher gender interactions on learning outcomes. Evidence from Sub-Saharan Africa *

Maria Kuecken†   Marie-Anne Valfort‡

October 29, 2012

Abstract

Using a within-student variation across subjects, we examine how student-teacher gender interactions impact primary school learning outcomes in mathematics and reading in eleven Sub-Saharan African countries. Our findings reveal that both boys and girls perform better with a female teacher in reading rather than with a male teacher. The reverse is true for math—male teachers are better for all students. These findings suggest that the traditional academic stereotype according to which "males are good at math and females are good at reading" plays a dominant role in explaining the impact of student-teacher gender interactions on academic achievement.

Keywords: Student-teacher gender interactions, Educational quality, Sub-Saharan Africa, Primary education

JEL: A12, I21, J16, N37

*The research presented in this paper is not the result of any for-pay consulting relationship.
†Paris School of Economics - Paris 1 Panthéon Sorbonne University; maria.kuecken@malix.univ-paris1.fr.
‡Paris School of Economics - Paris 1 Panthéon Sorbonne University; marie-anne.valfort@univ-paris1.fr.

1 Introduction

In developing countries, a once-narrow focus on improved access to schooling has given way to more nuanced concerns about student learning outcomes. Recent learning assessments across Sub-Saharan Africa and South Asia reveal that large numbers of students leave primary school without acquiring even the most basic competencies in reading and mathematics (Pratham (2005), Uwezo (2010), Brookings Institution (2011), UNESCO (2011)).

In this paper, we analyze the impact of student-teacher gender interactions on primary school children's performance in Sub-Saharan Africa. Sub-Saharan Africa has received little attention on this issue relative to developed countries, while this continent maintains one of the lowest levels of student achievement worldwide. Investigating whether specific student-teacher gender interactions could improve such achievement is therefore a critical step toward sound educational policies in this part of the world. This is especially true given that, despite a lack of evidence-based consensus, existing policies already seek to promote same-gender student-teacher pairings as a method to improve the enrollment, retention and achievement of girls who are known to lag behind relative to their male peers (UNESCO (2006), Lloyd (2009), UNESCO (2010)).

But whether it is better for a student to be matched with a teacher of the same gender or not is difficult to foretell ex-ante. Two classes of explanations for the educational relevance of a teacher's gender (possibly depending on the subject taught) have been identified by the literature.

The first class involves how students respond to a teacher's gender, encompassing two effects: the role model effect and the stereotype effect. The role model effect induces students to be more engaged and responsive with teachers of their own gender. It therefore predicts that same-gender student-teacher matching should positively impact both female and male students' performance. The stereotype effect refers to the traditional academic stereotype according to which "males are good at math and females are good at reading".[1] This effect influences student expectations about teachers' proficiencies, according to which female teachers will be worse at teaching math and male teachers will be worse at teaching reading.

[1] Fryer and Levitt (2010) document the gender gap in math in which males outperform females. Other works like Guiso, Monte, Sapienza, and Zingales (2008) also document a gender gap in reading in which females outperform males, although the evidence for this gap is weaker in developing countries (Alderman, Behrman, Ross, and Sabot (1996), Brookings Institution (2011)). In our Sub-Saharan African data, we do observe a gender gap in math (with male math teachers and male students performing significantly better at standardized math tests than their female counterparts) as well as a gender gap in reading (with female reading teachers and female students performing better at standardized reading tests than their male counterparts). However the significance of the latter gap, both statistically and in terms of magnitude, is weaker.

Its prediction is thus that students (both male and female) should perform better with a female teacher in reading than with a male teacher in reading, while students (both male and female) should perform better with a male teacher in math.[2] This outcome is likely to be reinforced by the fact that the stereotype effect also influences teacher self-expectations, with male math teachers feeling more confident in their teaching ability than female math teachers, and female reading teachers feeling more confident in their teaching ability than male reading teachers.[3]

The second class of explanations for the educational relevance of a teacher's gender involves how teachers respond to student gender, provided this response is specific to teacher gender.[4] Overall, educational studies do not support a gender bias effect whereby teachers would give higher (or lower) assessments to children of their own gender (see Ouazad (2008). However, a sexual harassment effect whereby male teachers explicitly target female students or implicitly tolerate harassment perpetuated by male students on female peers is prevalent, though under-reported, in developing countries (Bott, Morisson, and Ellsberg (2005), Dunne (2007), Bhana (2012), Chisamya, DeJaeghere, Kendall, and Khan (2012)). The sexual harassment effect therefore predicts that female students should perform better with a female teacher, regardless of the subject taught.[5]

Measuring the impact of student-teacher gender interactions on students' academic performance (and hence the overall outcome of the aforementioned role model, stereotype, and sexual harassment effects) is challenging. Cross-sectional evaluations of student-teacher gender interactions are likely to be biased by the non-random assignment of students to teachers. As an illustration, suppose that such evaluations point to a positive relationship between same-gender student-teacher matching and students' academic performance. This evidence could be simply driven by the fact that female students with an unobserved propensity for

[2]Indeed, Lam, Tse, Lam, and Loh (2009) show that both male and female students from Hong Kong perform better with female teachers in reading. Echoing this result, Chudgar and Sankar (2008) show that all students perform better with a female reading teacher in five Indian states while, in Pakistan, Warwick and Haroona (1994) report that all students perform better with male teachers in math.

[3]As Beilock, Gunderson, Ramirez and Levin (2010) highlight, math anxiety is widespread among elementary school female teachers, leading to poorer math achievement amongst their students (as compared to those students matched with male math teachers).

[4]The stereotype effect therefore does not belong to this second class of explanations. Rather, it implies that *both* male and female teachers believe that boys have a higher propensity for achievement in math while girls have a higher propensity for achievement in reading. As such expectations are often self-fulfilling (or "Pygmalian", after Rosenthal and Jacobson (1968)), they can lead boys to perform better than girls in math and girls to perform better in reading, *regardless* of the teacher's gender (see Spencer, Steele and Quinn (1999) for an illustration of this mechanism in math).

[5]For example, the Brookings Institution (2011) provides evidence that a proliferation of female teachers can be an effective method in improving the safety of girls by preventing them from becoming targets of discrimination and abuse.

achievement are more likely to be matched with female teachers (and *vice versa* for male students), not at all by the positive impact of matching a student to a same-gender teacher.

Our paper addresses this potential bias by evaluating the impact of a teacher's gender on primary school children in Sub-Saharan Africa in specifications that control for student fixed effects. To do so, we rely on the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ II) survey, completed in 2005.[6] SACMEQ administers standardized reading and mathematics examinations to both students and teachers in order to compare achievement in the final year of primary school across member countries.[7] Because, for each student, SACMEQ reports pairs of test scores in both math and reading, we are able to exploit these matched pairs by running a within-student analysis that controls for subject fixed effects as well as student fixed effects that are constant across subjects.

To be sure, a student fixed effect approach eliminates some of the potential bias. However, three main sources of bias could still compromise the internal validity of this method. First, the within-student analysis does not allow us to treat the bias induced by the non-random assignment of students to teachers if this non-random assignment concerns students with subject-specific propensities for achievement. However, we expect this bias to be weak given that our data reveal a very strong correlation (of 76%) between students' achievements in both reading and math. Put differently, students' propensities for achievement appear to be constant across subjects rather than subject-specific. Even so, we attempt to alleviate any potential for this bias by controlling for a critical characteristic for achievement that is both student- and subject-specific—students' access to a textbook in math and reading. Second, the within-student analysis does not allow us to address the bias induced by the non-random assignment of teachers to students. Suppose our analysis again points to a positive relationship between same-gender student-teacher matching and students' academic performance. This evidence could very simply be driven by the fact that female teachers with higher competencies and/or better access to teaching resources are more likely to be matched with female students (and *vice versa* for male teachers). To decrease the likelihood of observing this bias, we control for a rich set of teacher characteristics (their age, qualifications, standardized test scores, dedication to their students, access to teaching resources, etc.). Third, by using the within-subject analysis, we can treat the bias induced by the non-random assignment of students to teachers, provided that the correlation between them

[6]We opt to use only SACMEQ II as it reports more recent data than SACMEQ I as well as a wider variety of teacher information than SACMEQ III.

[7]We study Botswana, Lesotho, Kenya, Malawi, Mozambique, Namibia, Seychelles, Swaziland, Tanzania, Uganda, and Zambia. Our analysis excludes Mauritius and South Africa as neither country reports test scores for teachers, a crucial control variable in our analysis.

is not close to perfect. We rule out the possibility of a quasi-perfect correlation between students' propensities for achievement and their assignment to a teacher of a specific gender in the robustness checks.

This paper improves upon the literature regarding the impact of student-teacher gender interactions on learning outcomes by being the first to extend this type of student fixed effect approach (which has been recently adopted in many contributions dealing with developed countries[8]) to a wide range of Sub-Saharan African countries.[9] Unsurprisingly, our results reveal a larger impact of student-teacher gender interactions on student test scores than that found in developed countries.[10] Developing countries are indeed characterized by higher levels of gender inequality[11] which are likely to activate the three effects (role model effect, stereotype effect,[12] and sexual harassment effect) through which student-teacher gender interactions are expected to impact student test scores. More precisely, our findings reveal that both boys and girls perform better with female rather than male teachers in reading. The reverse occurs for math, in which both boys and girls perform better with male teachers. These findings suggest that the traditional academic stereotype according to which "males are good at math and females are good at reading" plays a dominant role in explaining the impact of student-teacher gender interactions on academic achievements.

The paper proceeds as follows: In Section 2, we introduce a conceptual framework that stresses the expected impact of the role model, stereotype and sexual harassment effects on student performance. We describe our data in Section 3. In Section 4, we present the estimation strategy. Section 5 presents the results, and Section 6 provides robustness checks. Finally, Section 7 summarizes our results, discusses their policy implications, and highlights

[8]See Ammermuller and Dolton (2006), Dee (2007), Ouazad (2008), Holmlund and Sund (2008), Neugebauer, Helbig, and Landmann (2011) and Cho (2012).

[9]Aslam and Kingdon (2011) extend this approach to the South Asian context by relying on Pakistani data. Yet, their focus is more on the impact of teaching practices than on the impact of student-teacher gender interactions. See also Warwick and Haroona (1994), Michaelowa (2001), UNESCO (2005), Asadullah, Chaudhury, and Dar (2007) and Chudgar and Sankar (2008) for complementary evidence on the impact of student-teacher gender interactions in the context of developing countries which does not rely on a within-student analysis.

[10]With the exception of Dee (2007) and Carrell, Page, and West (2010) who rely on data from the U.S. Air Force Academy where students are randomly assigned to professors, studies based on developed countries that aim to treat potential estimation biases find no impact of student-teacher gender interactions on student achievement.

[11]The World Development Report on Gender Equality and Development (2012) provides comprehensive evidence from a substantial number of studies that poor countries, and specifically the poorest within these countries, fare the worst in terms of gender equality across dimensions such as endowments, economic opportunities, and agency.

[12]Dickerson, McIntosh and Valente (2012), for example, show that girls' poor performance in math can be explained in large part by fertility and sexual status norms. Studies by Hyde and Mertz (2009) and Pope and Snydor (2010) also find a relation between gender inequality and female mathematics performance.

avenues for future research.

2 Conceptual framework

We consider the following simple function for the academic performance of student i in subject j ($j = \{\text{math}, \text{reading}\}$):

$$P_{ij} = P_{ij}(SG_{ij}, T_{ij}, \mathbf{W_j}, F_i). \tag{1}$$

Equation (1) indicates that the academic performance of student i in subject j is determined by student-teacher gender interactions (SG_{ij}), a dummy that is equal to 1 if the student and his/her teacher in subject j share the same gender and 0 otherwise. Moreover, Equation (1) assumes that P_{ij} depends on a subject-specific student characteristic (T_{ij}) that reports whether student i has access to a textbook in subject j, as well as on a vector of subject characteristics ($\mathbf{W_j}$) which contains information on the traits of the teacher in subject j (including the classroom resources to which the teacher in subject j has access). Finally, we include an individual fixed effect denoted by F_i.

We assume that P_{ij} is increasing with respect to SG_{ij} and that SG_{ij} is itself a function of three possible effects: a role model effect (Rol), a stereotype effect (Ste) and a sexual harassment effect (Sex), such that:

$$SG_{ij} = \begin{cases} SG_{ij} = SG_{ij}(\underset{+}{Rol}, \underset{+}{Ste}) & \text{if } i \text{ is male and } j = \text{math} \\ SG_{ij} = SG_{ij}(\underset{+}{Rol}, \underset{-}{Ste}) & \text{if } i \text{ is male and } j = \text{reading} \\ SG_{ij} = SG_{ij}(\underset{+}{Rol}, \underset{-}{Ste}, \underset{+}{Sex}) & \text{if } i \text{ is female and } j = \text{math} \\ SG_{ij} = SG_{ij}(\underset{+}{Rol}, \underset{+}{Ste}, \underset{+}{Sex}) & \text{if } i \text{ is female and } j = \text{reading} \end{cases}$$

These relations exhibit the potential avenues for the role model, stereotype, and sexual harassment effects that might be present within student-teacher gender interactions. Taken alone, we expect that the presence of a role model effect from having a same-gender teacher (relative to an opposite-gender teacher) will have a positive influence on testing for both boys and girls. A presence of the stereotype effect would suggest that boys will have higher achievement in math while girls will have higher achievement in reading, regardless of teacher gender. Finally we consider that, for girls, a third potential factor is present in the student-teacher dynamic—sexual harassment when assigned to a male teacher. Assignment to a same-gender teacher should circumvent the detrimental effects of sexual harassment and, thus, girls can be expected to perform better with same-gender teachers. As we have no detailed information on these factors, we cannot test each of them explicitly and instead speculate as to how they might interact to produce certain outcomes.

Thus, this simple conceptual framework predicts that a male student should always perform better with a male teacher in math while a female student should always perform better with a female teacher in reading. Regarding the impact of same-gender teacher matching for male students in reading and for female students in math, four empirical results could emerge: First, we could observe that this same-gender student-teacher pairing yields higher academic achievement for both male and female students. This would suggest that the role model effect dominates the stereotype effect (this effect being further reinforced by the existence of sexual harassment in the case of female students). Second, our results may reveal that this same-gender student-teacher pairing yields lower levels of academic achievement for both male and female students. This would indicate that the stereotype effect dominates the role model effect (as well as the sexual harassment effect in the case of female students). Third, we could find that this same-gender student-teacher pairing has no impact on male students while it yields higher academic achievement for female students. A result of this nature would suggest that the role model and stereotype effects compensate one another, thereby leading the sexual harassment effect to play a dominant role since, in our case, it is relevant to female students only. Finally, our results could show that this same-gender student-teacher pairing produces higher academic achievement among male students while it has no impact on the achievement of female students. This would suggest that the stereotype effect dominates the role model effect as long as it is not challenged by the influence of sexual harassment (if one does not focus on female student-male teachers pairings), in which case the stereotype and sexual harassment effects compensate one another. We put these theoretical mechanisms to the test in subsequent sections.

3 Data

Our educational data derive from the second round of a survey completed in 2005 by the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ). The SACMEQ II survey measures primary school educational outcomes across all member countries in a comparable way. More specifically, SACMEQ collects information about student and teacher subject knowledge via mathematics and reading tests as well as a wide range of school characteristics through student, teacher, and school director surveys. We cover eleven countries in our analysis, including Botswana, Kenya, Lesotho, Malawi, Mozambique, Namibia, Tanzania (mainland and Zanzibar), Seychelles, Swaziland, Uganda, and Zambia. We are forced to exclude Mauritius and South Africa as neither country reports test scores for teachers, a crucial control variable in our analysis.

3.1 Student performance

The SACMEQ dataset includes student scores on standardized reading and mathematics examinations which we utilize as our dependent variable to measure student achievement. Summary statistics for all variables are reported in Table 1, with test scores in Panel A.

3.2 Controls varying at the student level

Due to the fact that we run a within-student analysis, we need only incorporate observable characteristics that could influence reading and mathematics scores and vary at the student level. More precisely, traits that are both student- and subject-specific as well as traits that are both teacher- and subject-specific are instrumental to our estimation strategy. At the student/subject-specific level, we control for whether or not a student has access to a textbook, as defined by ownership or sharing. This variable is found in Panel B of Table 1.

At the teacher/subject-specific level, we account for teacher gender (using an indicator for females), age, and highest level of academic qualification obtained (with dummy variables for primary, junior secondary, senior secondary, and A-level/tertiary). We also capture elements of teacher competency, teaching style, and preferences. As teachers take standardized tests in their subject areas, we add teacher test scores as a control for their competency in their taught subject (with maximum scores of 41 and 47 respectively). We also account for the frequency with which teachers correct homework (never, sometimes, always), importance they assign to encouraging their students (not importance, of some importance, very important), and frequency with which they assess their students (no test, once per year, once per term, 2-3 times per term, 2-3 times per month, once or more per week). Additionally, we include an average of dummy variables for the presence of specific classroom resources (such as writing board, chalk, wall chart, cupboard or locker, one or more bookshelves, classroom library or book corner, teacher table, and teacher chair). Panels D and E of Table 1 present these teacher/subject-specific characteristics.

3.3 Controls varying at the regional level

Further analyses to test the robustness of our results lead us to a within-region approach. We must therefore account for additional characteristics that vary by region. In addition to our original controls, we include characteristics of students (which are not subject-specific), classes, and schools. At the student level, we control for age as well as socioeconomic status with a proxy, employing an average of 14 home possessions[13] present in the student's household. This information can be found in Panel C of Table 1. At the classroom level, in Panel F of Table 1, we control for class characteristics such as size and ratio of girls within each class—the average class holds roughly 43 students, with girls comprising only about 10% of a class on average.

Finally, as shown in Panel G of Table 1, we consider school-specific traits. These include the frequency of teacher misbehavior (an average of eight misbehavior types such as absenteeism, tardiness, and using abusive language), student misbehavior that impacts own achievement (an average of seven misbehavior types such as tardiness, absenteeism, and skipping class) as well as the achievement of others (an average of three misbehavior types including bullying, sexually harassing, and fighting other students), and community involvement (an average of fourteen support activities offered by the community). We also include variables for school condition (with an indicator if the school is in good condition or needs only minor repair), resources (an average of whether the school possesses 23 resources such as a library, staff room, first aid kit, etc.), and school location (using an indicator if the school is in an urban area). Finally, we consider the school director's sex (using an indicator for females), age, years of experience, and dummy variables for the highest level of academic qualification attained.

4 Empirical strategy

Following studies such as Dee (2007), Aslam and Kingdon (2011), and Cho (2012), we run a within-student analysis across subjects using student fixed effects. Due to the unique matched pair nature of our data, with two test scores per student, estimating a fixed effects analysis is precisely equivalent to a first-difference analysis.[14] The utility of this strategy is such that unobserved student characteristics can be differenced out and, as a result,

[13]These include a newspaper, magazine, radio, television, VCR, cassette, telephone, refrigerator, car, motorcycle, bicycle, water, electricity, table.

[14]This approach is itself inherited from Ashenfelter and Kruger (1994), Ashenfelter and Rouse (1998), and Rouse (1999).

estimations are not biased by these unobservables.

Yet, this strategy can fall prey to three potential biases. First, the within-student analysis does not allow us to treat the bias induced by the non-random assignment of students to teachers if this non-random assignment concerns students with subject-specific propensities for achievement. However, given that our data reveal a very strong correlation (of 76%) between students' achievements in both subjects, we expect this bias to be weak. In other words, students' propensities for achievement seem to be constant across subjects rather than subject-specific. Regardless, in this section we attempt to alleviate any potential for this bias by controlling for a critical characteristic for achievement that is both student- and subject-specific—a student's access to a textbook in math and reading. Second, the within-student analysis does not allow us to address the bias induced by the non-random assignment of teachers to students. To reduce the likelihood of observing this bias, we also control for a rich set of teacher characteristics (age, qualifications, standardized test scores, dedication to their students, access to teaching resources, etc.). Third, the within-subject analysis treats the bias resulting from the non-random assignment of students to teachers, provided that the correlation between them is not close to perfect. In the robustness checks, we rule out the possibility of a quasi-perfect correlation between students' propensities for achievement and their assignment to a teacher of a specific gender.

We make use of the following specification, estimated separately for male and female students with an OLS procedure:

$$P_{ij} = a_i + b.SG_{ij} + c.SG^{*}_{ij}MATH + d.MATH + e.T_{ij} + \mathbf{W'_j}.\mathbf{f} + \epsilon_{ij} \tag{2}$$

P_{ij} represents the test score for student i in subject j. Student fixed effects, captured by a_i, account for unobserved student traits. SG_{ij} is a dummy that takes the value of 1 if the teacher is of the same gender as the student and 0 otherwise. $MATH$ is a dummy that takes the value of 1 if the subject tested is math and 0 if reading. In keeping with the idea that gender-specific effects may vary by subject, we interact SG_{ij} with $MATH$. As a consequence, the coefficient b associated to SG_{ij} stands for the impact that a reading teacher of the same gender has on the reading performance of student i (as opposed to the impact of having a reading teacher of the opposite gender). In the same vein, the sum of coefficients b and c represents the impact of a same-gender math teacher on the math performance of student i (as opposed to the impact of having a math teacher of the opposite gender). We include a variable T_{ij} representing a trait that is simultaneously specific to both students and subjects (access to a math or reading textbook) as well as a vector of teacher traits ($\mathbf{W'_j}$)

to control for observable characteristics in the learning environment. Finally, we include the mean-zero error term (ϵ_{ij}) and cluster standard errors at the school level.

5 Results

We present OLS estimations from Equation (2) in Table 2 with results listed separately in odd and even columns for boys and girls, respectively. Controls are introduced step-wise, with Wald tests for gender interactions and their associated p-values located at the bottom of the table. Regarding boys, the first Wald test at the bottom of Table 2 in Column 1 reveals that they perform better with female reading teachers than with male reading teachers. However, the subsequent Wald test indicates that they have higher achievement in math when matched with a male math teacher. The same is true for girls—they have higher achievement with female reading teachers and male math teachers, as evidenced by the first and second Wald tests in Column 2.

More specifically, being matched to a female reading teacher increases boys' test scores by 0.09 standard deviations and girls' test scores by 0.14 standard deviations. In math, assignment to a male teacher increases boys' scores by 0.09 standard deviations and girls' scores by 0.06 standard deviations.[15] When compared alongside various educational interventions (as described by Kremer and Holla (2009)), the impacts of student-teacher gender interactions on test scores are more substantial than those associated to giving prizes to teachers/school committees and providing inputs such as flipcharts or textbooks. Moreover, the comparatively large impact of a girl being paired with a female reading teacher has an effect on par with (short-run) interventions such as linking teacher pay to test scores and tracking.

In general, these results suggest that the stereotype factor is at work, according to which female teachers are assumed to be better at reading instruction and male teachers to be better at math instruction. The fact that opposite-gender student-teacher matching is more beneficial for boys in reading and girls in math shows that, even if the other role model or sexual harassment effects are present, they are unlikely to be driving our results. For example, if the role model effect were dominant, we would see a positive impact of same-

[15]For reading, we calculate these values as the estimated coefficient of the dummy for a same-gender teacher (SG_{ij}) divided by the standard deviation of reading scores (for boys and girls separately). For math, we calculate these values as the sum of the estimated coefficients of the dummy for a same-gender teacher (SG_{ij}) and the interaction term $SG_{ij}^{*}MATH$ divided by the standard deviation of math scores (for boys and girls separately). Note that in both cases we rely on the most complete specification—that is, on Columns 5 and 6 of Table 2.

gender student-teacher matching for boys and girls in both subjects—this is not the case, as same-gender student-teacher matching is detrimental for boys in reading as well as for girls in math. Moreover, though the negative effects of sexual harassment may exist for female students assigned to male math teachers, we do not observe a negative impact overall in their math achievement. Rather, the stereotype effect seems to play a dominant role in our results.

6 Robustness

In keeping with Dee (2007), it may be more likely for teachers of a certain gender in math or reading to be assigned to students with propensities for higher or lower achievement. For example, our result that both boys and girls perform better with female reading teachers may be a result of specific student-teacher gender interactions or, alternatively, the fact that female reading teachers are not randomly assigned to students. In other words, it could be that female reading teachers are assigned to students with higher propensities for achievement regardless of the subject. If this latter explanation were true, it would explain our results without the dynamic of gender interactions.

Controlling for student fixed effects is a useful strategy for it allows us to limit the possibility that our results are driven by a non-random assignment of students to teachers. However, if the correlation between propensity for achievement and assignment is nearly perfect, our approach would no longer eliminate this potential bias. Therefore, our robustness checks aim to rule out the existence of a quasi-perfect correlation between students' potential and their teacher assignment.

One way to rule out this potential for non-random assignment is therefore to test our two main results by observing (1) if students who have a female reading teacher perform better overall (both in reading and math) than those having a male reading teacher and (2) if students who have a male math teacher perform better overall than those having a female math teacher. If we observe that either (1) or (2) is true, then a high correlation potentially exists between students' propensity for achievement and their teacher assignment that a within-student analysis may be unable to treat.

6.1 Are better students assigned to female reading teachers?

To test whether or not students assigned to a female reading teacher are higher achievers overall, we introduce into our analysis a dummy variable to represent assignment to a female reading teacher (FR). As this variable is defined at the student level, we can no longer rely on a within-student analysis as in the previous section. Moreover, given that there are under two teachers per subject, per school on average, a within-school analysis cannot suffice either. We therefore turn to a within-region analysis[16] in which standard errors are clustered at the school level and regional fixed effects account for unobserved regional traits. This strategy is presented in Equation (3):

$$P_{ij} = a + b.FR_i + c.FR_i^{*}MATH + d.MATH + e.T_{ij} + \mathbf{S_i'}.\mathbf{f} + \mathbf{W_j'}.\mathbf{g} + \mathbf{X_k'}.\mathbf{h} + \mathbf{Z_m'}.\mathbf{i} + \epsilon_{ij} \quad (3)$$

As before, P_{ij} represents the test score for student i in subject j. The coefficient b represents the impact of a female reading teacher on student i's reading performance. For both boys and girls, we know this effect should be positive and significant. However, we do not yet know the impact of having a female reading teacher on the performance of student i in math, represented by the sum of coefficients b and c. Moreover, because we run a within-region analysis, we must control for factors that vary at the regional level, initially reported in Table 1. We therefore incorporate student and subject (T_{ij}), student ($\mathbf{S_i'}$), teacher ($\mathbf{W_j'}$), classroom ($\mathbf{X_k'}$), and school ($\mathbf{Z_m'}$) characteristics.

We run the check separately for boys and girls. Results are shown in separate columns of Table 3. The first Wald test at the bottom of Table 3 in Column 1 confirms that boys perform better in reading with a female reading teacher. However, the second Wald test reveals that being assigned to a female reading teacher has no impact on boys' math performance. Similarly, the first Wald test at the bottom of Table 3 in Column 2 shows that girls also perform better with a female reading teacher while the second Wald test highlights that being assigned to a female reading teacher has no impact on girls' performance in math. These outcomes rule out the possibility of a quasi-perfect correlation between a student's propensity for achievement and their assignment to a female reading teacher and, thus, give credence to the role of gender interactions in reading performance.

6.2 Are better students assigned to male math teachers?

We then test the possibility that students who have a male math teacher perform better in both subjects, and hence have a higher propensity for achievement, than those with a female math teacher. To do so, we substitute the variable FR (representing assignment to

[16]There are 95 regions in our dataset.

a female reading teacher) in Equation (3) with a dummy variable that captures whether or not a student has a male math teacher.

Results are presented in Table 4, with separate columns for boys and girls. As before, the second Wald test at the bottom of Table 4 affirms that both boys and girls do perform better in math with a male math teacher instead of a female math teacher. However, being assigned to a male math teacher is actually detrimental to boys' achievement in reading as shown by the first Wald test of Column 1. For girls, the first Wald test of Column 2 shows that this matching has no impact on their reading performance. These results therefore also reject the existence of a quasi-perfect correlation between a student's ability and their assignment to a male math teacher. Taken with the results of the previous section, both checks suggest that our results on the impact of student-teacher gender interactions on students' academic achievements are indeed driven by student-teacher gender interactions.

7 Conclusion

On matched pair test score data of student math and reading achievement, we rely on a within-student approach to estimate the impact of student-teacher gender matching on primary school learning outcomes in eleven countries. Our results reveal that both boys and girls obtain higher test scores when assigned to female teachers in reading and male teachers in math. These findings suggest that the traditional academic stereotype according to which "males are good at math and females are good at reading" plays a larger role in explaining the impact of student-teacher gender interactions on academic achievement than the role model or sexual harassment effects.

In the short-run, these results point to recommendations that are at odds with existing policies for improving academic achievement. Despite a lack of consistent evidence, current policies related to student-teacher gender matching assume that the main barrier to primary school students' (and specifically girls') achievement is sexual discrimination/harassment perpetuated by both male teachers and students. Such policies therefore seek to promote same-gender student-teacher matching as a method to improve enrollment, retention and achievement among girls. In contrast, our findings point to the benefits of matching both boys and girls with male teachers in math and female teachers in reading.

However, generalizing such a policy would evidently be detrimental in the long-run as it would reinforce the stereotype according to which reading and math are the domains of females and males respectively. The goal of policy-makers should therefore be to destroy any channel through which student-teacher gender interactions could influence student achievement since all channels (stereotype, role model and sexual harassment effects) are ultimately tied to some gender-based prejudice. Regarding the stereotype effect, one should attempt to replace stereotypes about gender-specific achievement with the notion that students and teachers of any gender are capable of performing well in both reading and math.[17] Eliminating the role model effect would require overcoming students' tendencies to be more engaged and responsive with teachers of their own gender. Finally, to prevent sexual harassment, policies would need to initiate teacher training in order to make all teachers (especially males) aware of proper conduct as well as promote effective enforcement procedures in schools. Achieving these three goals constitutes a critical avenue for future research.

References

[1] Alderman, Harold, Jere R. Behrman, David R. Ross, and Richard Sabot. 1996. "Decomposing the gender gap in cognitive skills in a poor rural economy." *Journal of Human Resources*, 31: 229-254

[2] Ammermuller, Andreas and Peter Dolton. 2006. "Pupil-teacher gender interaction effects on scholastic outcomes in England and the USA." ZEW Discussion Paper No. 06-060. Centre for European Economic Research.

[3] Aslam, Monazza and Geeta Kingdon. 2011. "What can teachers do to raise pupil achievement?" *Economics of Education Review*, 30: 559-574.

[4] Asadullah, Mohammad N., Nazmul Chaudhury, and Amit Dar. 2007. "Student Achievement Conditioned upon School Selection: Religious and Secular Secondary School Quality in Bangladesh." *Economics of Education Review*, 26(6): 648-659.

[5] Beilock, Sian L., Elizabeth A. Gunderson, Gerardo Ramirez, and Susan C. Levine. 2010. "Female teachers'math anxiety affects girls' math achievement." *Proceedings of the National Academy of Sciences of the United States of America*, 107(5): 1860-1863.

[6] Bhana, Deevia. 2012. "Girls are not free—In and out of the South African school." *International Journal of Educational Development*, 32(2): 352-358.

[17] Guiso, Monte, Sapienza, and Zingales (2008), Fryer and Levitt (2010), Cvencek, Meltzoff, and Greenwald (2011), and Dickerson, McIntosh, and Valente (2012) suggest that there is a cultural justification for gender gaps in math rather than a genetic predisposition for achievement according to gender.

[7] Bott, Sarah, Andrew Morrison, and Mary Ellsberg. 2005. "Preventing and responding to gender-based violence in middle and low-income countries: a global review and analysis." World Bank Policy Research Working Paper (WPS 3618).

[8] The Brookings Institution. 2011. *A Global Compact on Learning: Taking Action on Education in Developing Countries.*

[9] Carrell Scott. E, Marianne E. Page, and James E. West. 2010. "Sex and science: How professor gender perpetuates the gender gap." *The Quarterly Journal of Economics*, 125(3): 1101-1144.

[10] Chisamya, Grace, Joan DeJaeghere, Nancy Kendall, and Marufa Aziz Khan. 2012. "Gender and Education for All: Progress and problems in achieving gender equity." *International Journal of Educational Development*, 32(6): 743-755.

[11] Cho, Insook. 2012. "The effect of teacher-student gender matching: Evidence from OECD countries." *Economics of Education Review*, 31: 54-67.

[12] Chudgar, Amita and Vyjayanthi Sankar. 2008. "The relationship between teacher gender and student achievement: evidence from five Indian States." *Compare: A Journal of Comparative and International Education*, 38(5): 627-642.

[13] Cvencek, Dario, Andrew N. Meltzoff, and Anthony G. Greenwald. 2011. "Math-Gender Stereotypes in Elementary School Children." *Child Development*, 82(3): 766-779.

[14] Dee, Thomas S. 2007. "Teachers and the gender gaps in student achievement." *Journal of Human Resources*, 42: 525-554.

[15] Dei, George J. Sefa. 2004. "Dealing with difference: ethnicity and gender in the context of schooling in Ghana." *International Journal of Educational Development*, 24: 343-359.

[16] Dickerson, Andy, Steven McIntosh, and Christine Valente. 2012. "Do the Maths: An analysis of the gender gap in mathematics in Africa." *Working paper.*

[17] Dunne, Máiréad. 2007. "Gender, sexuality and schooling: Everyday life in junior secondary schools in Botswana and Ghana." *International Journal of Educational Development*, 27(5): 499-511.

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

[19] Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales. 2008. "Culture, Gender, and Math." *Science*, 320: 1164-1165.

[20] Holmlund, Helena and Krister Sund. 2008. "Is the gender gap in student performance affected by the sex of the teacher?" *Labour Economics*, 15(1): 37-53.

[21] Hyde, Janet S. and Janet E. Mertz. 2009. "Gender, culture, and mathematics performance." *Proceedings of the National Academy of Sciences of the United States of America*, 106(22): 8801-8807.

[22] Kremer, Michael and Alaka Holla. 2009. "Improving education in the developing world: what have we learned from randomized evaluations?" *Annual Review of Economics*, 1: 513-542.

[23] Lam, Y.H. Raymond, Shek Kam Tse, Joseph W.I. Lam, and Elizabeth K.Y. Loh. 2009. "Does the gender of the teacher matter in the teacher of reading literacy? Teacher gender and pupil attainment in reading literacy in Hong Kong." *Teaching and Teacher Education*, 26: 754-759.

[24] Lloyd, Cynthia B. 2009. *New Lessons: The Power of Educating Adolescent Girls.* New York: Population Council.

[25] Michaelowa, Katharina. 2001. "Primary education quality in Francophone sub-Saharan Africa:Determinants of learning achievement and efficiency considerations." *World Development*, 29(10): 1699-1716.

[26] Neugebauer, Martin, Marcel Helbig, and Andreas Landmann. 2011. "Unmasking the myth of the same-sex teacher advantage." *European Sociological Review*, 13(2): 125-145.

[27] Nguyen, Phuong L. 2006. "Effects of social class and school conditions on educational enrollment and achievement of boys and girls in rural Viet Nam." *International Journal of Educational Research*, 45(3): 153-175.

[28] Ouazad, Amine. 2008. "Assessed by a teacher like me: Race, gender and subjective evaluations." CEE Discussion Papers 0098. Centre for the Economics of Education, London School of Economics and Political Science.

[29] Pope, Devin G. and Justin R. Snydor. 2010. "Geographic Variation in the Gender Differences in Test Scores." *Journal of Economic Perspectives*, 24(2): 95-108.

[30] Pratham. 2005. *ASER 2005: Annual Status of Education Report.* Pratham Resource Center Mumbai.

[31] Rosenthal, R. and Jacobson, L. 1968. *Pygmalion in the Classroom*, Holt, Rinehart and Winston, New York.

[32] Spencer, Steven J., Claude M. Steele, and Diane M. Quinn. 1999. "Stereotype threat and women's math performance." *Journal of Experimental Social Psychology*, 35(1): 4-28

[33] UNESCO. 2005. *EFA Global Monitoring Report 2005: The quality imperative.* Paris: UNESCO.

[34] UNESCO. 2006. *Advocacy brief: The impact of women teachers on girls' education.* Bangkok: UNESCO.

[35] UNESCO. 2010. *EFA Global Monitoring Report 2010: Reaching the Marginalized.* Paris: UNESCO.

[36] UNESCO. 2011. *EFA Global Monitoring Report 2011: The hidden crisis: Armed Conflict and education.* Paris: UNESCO.

[37] Uwezo Tanzania. 2010. *Are our children learning?* Annual Learning Assessment Report Tanzania.

[38] Warwick, Donald P. and Jatoi Haroona. 1994. "Teacher gender and student achievement in Pakistan." *Comparative Education Review*, 38(3): 377-399.

[39] World Development Report: Gender Equality and Development. 2012.

8 Tables

Table 1: Summary statistics

	Mean	Standard Deviation	Observations
Panel A: Dependent variable			
Student math score	494.04	92.26	36,829
Student reading score	498.36	96.58	37,062
Panel B: Student/subject characteristics			
Student has access to a math textbook	0.88	0.33	37,062
Student has access to a reading textbook	0.89	0.31	37,062
Panel C: Student characteristics			
Student gender (female)	0.49	0.50	37,058
Student age	18.60	1.89	37,062
Student home possessions	0.37	0.24	37,062
Panel D: Math teacher characteristics			
Gender (female)	0.46	0.50	2,644
Age	35.38	8.13	2,679
Qualification (primary)	0.11	0.31	2,679
Qualification (junior secondary)	0.21	0.41	2,679
Qualification (senior secondary)	0.50	0.50	2,679
Qualification (A-level/tertiary)	0.19	0.39	2,679
Teacher test score	25.75	6.84	2,625
Frequency of correcting homework	2.65	0.34	2,657
Importance of encouraging students	2.77	0.47	2,646
Frequency of assessing students	4.88	0.91	2,660
Average resources	0.62	0.26	2,659
Panel E: Reading teacher characteristics			
Gender (female)	0.53	0.50	2,713
Age	35.36	8.24	2,747
Qualification (primary)	0.12	0.32	2,747
Qualification (junior secondary)	0.20	0.40	2,747
Qualification (senior secondary)	0.47	0.50	2,747
Qualification (A-level/tertiary)	0.21	0.41	2,747
Teacher test score	30.75	5.66	2,729
Frequency of correcting homework	2.55	0.37	2,656
Importance of encouraging students	2.76	0.48	2,747
Frequency of assessing students	5.28	0.90	2,747
Average resources	0.62	0.25	2,731
Panel F: Class characteristics			
Class size	42.23	14.40	4,523
Class girl ratio	0.11	0.12	4,496
Panel G: School characteristics			
Average teacher misbehavior	1.44	0.34	1,959
Student misbehavior (own grades)	1.71	0.33	1,959
Student misbehavior (peers' grades)	1.76	0.44	1,959
Community involvement	0.38	0.24	1,959
School condition (good/minor repair)	0.50	0.50	1,959
School resources	0.35	0.20	1,959
School location (urban)	0.39	0.49	1,959
Director gender (female)	0.28	0.45	1,956
Director age	45.28	7.71	1,958
Director qualification (primary)	0.10	0.30	1,959
Director qualification (junior secondary)	0.15	0.36	1,959
Director qualification (senior secondary)	0.45	0.50	1,959
Director qualification (A-level/tertiary)	0.31	0.46	1,959
Director experience	21.47	7.96	1,959

Notes: In our dataset, there are 37,062 students, 2,679 math teachers, 2,747 reading teachers, 4,523 classes, and 1,959 schools.

Table 2: Student test scores and student-teacher gender interactions: OLS results

	Dep. var.: Student test scores					
	Boys	Girls	Boys	Girls	Boys	Girls
	(1)	(2)	(3)	(4)	(5)	(6)
Same gender	-8.890*** (2.743)	12.828*** (3.090)	-8.900*** (2.738)	12.820*** (3.094)	-8.272*** (2.844)	13.507*** (3.181)
Same gender*Math	18.058*** (2.699)	-19.689*** (3.084)	18.069*** (2.697)	-19.691*** (3.089)	16.687*** (2.718)	-18.982*** (3.187)
Math	-6.994*** (1.788)	-2.228 (1.887)	-7.034*** (1.790)	-2.186 (1.887)	-5.103** (2.249)	-1.732 (2.295)
Student has access to a textbook			-4.554 (3.024)	2.579 (3.487)	-2.101 (2.874)	2.918 (3.817)
Teacher age					0.122 (0.146)	0.282 (0.177)
Teacher qualification (junior secondary)					-4.625 (4.381)	-5.958 (4.881)
Teacher qualification (senior secondary)					-6.625 (4.449)	-4.361 (4.529)
Teacher qualification (A-level/tertiary)					-4.813 (4.945)	-2.849 (5.396)
Teacher test score					0.443** (0.178)	0.324 (0.200)
Frequency of correcting homework					2.972 (3.283)	1.195 (3.803)
Importance of encouraging students					1.204 (2.074)	0.545 (2.352)
Frequency of assessing students					-0.519 (1.029)	-0.438 (1.199)
Teacher's average access to resources					6.162 (7.396)	7.111 (8.309)
R^2	0.882	0.888	0.882	0.888	0.889	0.893
Observations	36,180	35,206	36,180	35,206	34,492	33,697
Student fixed effects	yes	yes	yes	yes	yes	yes
	Wald test p-values					
Same gender = 0	0.0012	0.0000	0.0012	0.0000	0.0037	0.0000
Same gender + Same gender*Math = 0	0.0009	0.0218	0.0009	0.0218	0.0044	0.0883

Notes: This table presents OLS estimates for boys and girls. Student test scores is the dependent variable. Controls are introduced step-wise, first with an interaction term for a same-gender teacher and a subject fixed effect, followed by student/subject-specific characteristics and, finally, teacher/subject-specific characteristics. The first Wald test reports the significance of the impact of a same-gender reading teacher on students' reading performance (as opposed to the impact of having a reading teacher of the opposite gender). The second Wald test reports the significance of the impact of a same-gender math teacher on students' math performance (as opposed to the impact of having a math teacher of the opposite gender). Standard errors are clustered at the school level, and student fixed effects are included in every column. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Table 3: Student test scores with female reading teachers: OLS results

	Dep. var.: Student test scores	
	Boys	Girls
	(1)	(2)
Female reading teacher	9.174*** (3.329)	10.598*** (3.168)
Female reading teacher*Math	-10.014*** (2.025)	-12.880*** (2.112)
Math	13.280*** (1.516)	0.212 (1.603)
Student has access to a textbook	9.634*** (1.858)	9.578*** (2.099)
Student age	-4.982*** (0.372)	-6.756*** (0.396)
Student home possession	42.783*** (4.030)	45.385*** (3.911)
Teacher gender (female)	0.771 (2.484)	-1.115 (2.272)
Teacher age	-0.010 (0.105)	0.080 (0.099)
Teacher qualification (junior secondary)	0.544 (2.978)	0.773 (2.827)
Teacher qualification (senior secondary)	-2.145 (2.996)	-0.901 (2.743)
Teacher qualification (A-level/tertiary)	5.643 (3.444)	8.502*** (3.171)
Teacher test score	1.085*** (0.148)	0.961*** (0.144)
Frequency of correcting homework	9.298*** (2.590)	6.542*** (2.468)
Importance of encouraging students	2.737* (1.593)	-0.246 (1.546)
Frequency of assessing students	1.506* (0.900)	0.299 (0.868)
Teacher's average access to resources	1.574 (4.491)	7.400* (4.269)
Class size	0.014 (0.095)	0.068 (0.103)
Class girl ratio	63.879*** (13.023)	67.245*** (13.120)
Average teacher misbehavior	0.222 (3.961)	-2.962 (3.928)
Student misbehavior (own grades)	-3.479 (4.218)	-1.424 (4.046)
Student misbehavior (peers' grades)	-1.704 (2.945)	-1.444 (3.008)
Community involvement	10.712 (6.750)	4.505 (5.937)
School condition	2.797 (2.093)	4.728** (1.884)
School resources	83.920*** (9.797)	82.999*** (8.535)
School location	11.340*** (2.351)	14.078*** (2.270)
Director gender (female)	1.758 (2.518)	3.677* (2.209)
Director age	0.120 (0.268)	0.241 (0.254)
Director qualification (junior secondary)	-1.098 (3.491)	-1.739 (3.162)
Director qualification (senior secondary)	1.133 (3.662)	0.162 (3.214)
Director qualification (A-level/Tertiary)	1.129 (3.924)	2.380 (3.417)
Director experience	-0.221 (0.264)	-0.247 (0.249)
R^2	0.357	0.389
Observations	33,023	31,587
Region fixed effects	yes	yes
	Wald test p-values	
Female reading teacher = 0	0.0059	0.0008
Female reading teacher + Female reading teacher*Math = 0	0.7442	0.3440

Notes: This table presents OLS estimates for boys and girls. Student test scores is the dependent variable. The first Wald test reports the significance of the impact of being assigned to a female reading teacher on students' reading performance (as compared to the impact of being assigned to a male reading teacher). The second Wald test reports the significance of the impact of being assigned to a female reading teacher on students' math performance (as compared to the impact of being assigned to a male reading teacher). Standard errors are clustered at the school level, and region fixed effects are included in every column. *, ** and *** indicate significance at the 10, 5 and 1% levels.

Table 4: Student test scores with male math teachers: OLS results

	Dep. var.: Student test scores	
	Boys	Girls
	(1)	(2)
Male math teacher	-6.585** (2.775)	-0.803 (2.661)
Male math teacher*Math	13.666*** (2.015)	13.289*** (2.084)
Math	1.028 (1.677)	-12.458*** (1.736)
Student has access to a textbook	8.982*** (1.879)	8.393*** (2.091)
Student age	-4.922*** (0.370)	-6.737*** (0.392)
Student home possession	42.356*** (3.989)	45.440*** (3.878)
Teacher gender (female)	7.746*** (2.354)	10.368*** (2.473)
Teacher age	-0.032 (0.106)	0.077 (0.099)
Teacher qualification (junior secondary)	-0.049 (2.936)	1.441 (2.825)
Teacher qualification (senior secondary)	-2.103 (2.882)	-0.272 (2.707)
Teacher qualification (A-level/tertiary)	5.937* (3.338)	8.695*** (3.130)
Teacher test score	1.062*** (0.152)	0.997*** (0.149)
Frequency of correcting homework	8.313*** (2.626)	4.819** (2.436)
Importance of encouraging students	3.178** (1.587)	-0.191 (1.531)
Frequency of assessing students	1.567* (0.895)	-0.020 (0.873)
Teacher's average access to resources	1.324 (4.580)	6.396 (4.277)
Class size	-0.002 (0.095)	0.042 (0.102)
Class girl ratio	64.011*** (12.973)	63.894*** (13.135)
Average teacher misbehavior	-0.135 (3.999)	-5.315 (3.713)
Student misbehavior (own grades)	-4.228 (4.181)	-1.614 (3.961)
Student misbehavior (peers' grades)	-2.122 (2.950)	-1.949 (2.983)
Community involvement	10.977 (6.894)	5.220 (5.964)
School condition	2.913 (2.108)	5.074*** (1.874)
School resources	82.342*** (9.907)	79.251*** (8.473)
School location	11.292*** (2.352)	14.238*** (2.271)
Director gender (female)	2.019 (2.526)	3.490 (2.207)
Director age	0.154 (0.268)	0.254 (0.254)
Director qualification (junior secondary)	-1.697 (3.532)	-2.469 (3.156)
Director qualification (senior secondary)	1.045 (3.701)	-0.177 (3.200)
Director qualification (A-level/Tertiary)	1.764 (3.981)	3.122 (3.414)
Director experience	-0.227 (0.265)	-0.218 (0.250)
R^2	0.361	0.393
Observations	32,761	31,334
Region fixed effects	yes	yes
	Wald test p-values	
Male math teacher = 0	0.0177	0.7628
Male math teacher + Male math teacher*Math = 0	0.0306	0.0001

Notes: This table presents OLS estimates for boys and girls. Student test scores is the dependent variable. The first Wald test reports the significance of the impact of being assigned to a male math teacher on students' reading performance (as compared to the impact of being assigned to a female math teacher). The second Wald test reports the significance of the impact of being assigned to a male math teacher on students' math performance (as compared to the impact of being assigned to a female math teacher). Standard errors are clustered at the school level, and region fixed effects are included in every column. *, ** and *** indicate significance at the 10, 5 and 1% levels.