

Africa's Skill Tragedy: Does Teachers' Lack of Knowledge Lead to Low Student Performance?^{*}

Jan Bietenbeck

Marc Piopiunik

Lund University and IZA

Ifo Institute and CESifo

Simon Wiederhold

KU Eichstätt-Ingolstadt, Ifo Institute, and CESifo[†]

We study the importance of teacher subject knowledge for student performance in Sub-Saharan Africa using unique international assessment data for sixth-grade students and their teachers. To circumvent bias due to unobserved student heterogeneity, we exploit variation within students across math and reading. Teacher subject knowledge has a modest impact on student performance. Exploiting huge cross-country differences in economic development, we find that teacher knowledge is effective only in more developed African countries. Results are robust to adding teacher fixed effects and accounting for potential sorting based on subject-specific factors.

JEL Code: I21, J24, O15

Keywords: teacher knowledge, student performance, Sub-Saharan Africa

*We thank Martin Abel, Manuel Arellano, Manuel Bagues, David Dorn, Joshua S. Goodman, Eric A. Hanushek, Ricardo Hausmann, Asim Khwaja, Alessandro Martinello, Frank Neffke, Randall Reback, Ludger Woessmann, and numerous seminar and conference audiences for helpful comments. We are grateful to SACMEQ for providing the data used in the empirical analysis in this paper. Bietenbeck acknowledges funding from the Spanish Ministry of Science and Innovation (BES-2011-050947). Wiederhold is grateful for the hospitality provided by the Center for International Development at Harvard University, especially to Ricardo Hausmann, Ljubica Nedelkoska, and Frank Neffke. Wiederhold also greatly appreciates the financial support provided by the Fritz Thyssen Foundation for financing the research stay at Harvard University.

[†]Piopiunik (corresponding author): Ifo Institute, Poschingerstraße 5, 81679 Munich, Germany; Email: piopiunik@ifo.de; Phone: +49 89 9224 1312. Bietenbeck: Lund University, P.O. Box 7082, 220 07 Lund, Sweden; Email: jan.bietenbeck@nek.lu.se. Wiederhold: Ifo Institute, Poschingerstraße 5, 81679 Munich, Germany; Email: wiederhold@ifo.de.

1 Introduction

Developing countries have made considerable progress in increasing school enrollment (Glewwe et al., 2013). However, children in these countries are still learning remarkably little in school. For example, in a large-scale assessment across Sub-Saharan Africa, sixth-grade students were asked to calculate the number of pages remaining in a 130-page book after the first 78 pages have been read. Only 30% were able to answer this question correctly. In comparison, two-thirds of *fourth*-grade students in OECD countries could answer this question.¹ The average performance of students in Sub-Saharan Africa is also dismal when compared to that of students in other comparable countries (for a comparison with students in India, see Hanushek and Woessmann, 2012). These are alarming findings for Sub-Saharan Africa since previous studies show that it is skills, and not the number of years spent in school, that drive economic growth (Hanushek and Woessmann, 2012).

While student performance is low across the entire region of Sub-Saharan Africa, there are substantial differences between countries. For example, correct-answer rates for the question described above range from 14% in Malawi to almost 50% in Kenya and Tanzania. This variation is unlikely to be explained by differences in school resources, given that the most convincing evidence from randomized interventions shows that resources have, at best, small effects on student performance (for a survey, see Murnane and Ganimian, 2014). In contrast, a growing literature documents the importance of teachers (e.g., Chetty, Friedman and Rockoff, 2014), suggesting that teacher quality plays a role in the observed cross-country differences in student performance.

We estimate the causal effect of one main dimension of teacher quality, teacher subject knowledge, on student performance in Sub-Saharan Africa. We use unique data that provide consistent measures of teacher subject knowledge and student performance for 13 countries.

¹These figures are based on data from the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) and the Third International Mathematics and Science Study (TIMSS), respectively. The question reads: “Tanya has read the first 78 pages of a book that is 130 pages long. Which number sentence could Tanya use to find the number of pages she must read to finish the book?” Students had to choose between the correct answer, $130 - 78 = X$, and three incorrect answers: $130 + 78 = X$, $X - 78 = 130$, and $130/78 = X$. Other comparable questions reveal similarly large performance gaps.

Exploiting math and reading measures, we identify the impact of teacher subject knowledge based only on differences within students between these two subjects. This eliminates any unobserved student heterogeneity that is constant across math and reading. In addition, we control for various subject-specific teacher characteristics and school resources.

Teacher subject knowledge has a positive and significant effect on student performance. Our within-student estimates indicate that increasing teacher subject knowledge by one standard deviation (SD) raises student performance by about 0.03 SD. Assuming that the variation in teacher effectiveness in Sub-Saharan Africa is similar to that in the United States, this implies that teacher subject knowledge explains about 20% of the variation in teachers' overall effectiveness.² Although the estimated teacher effect is rather modest, it is similar to the impact of other educational interventions, such as a 10% increase in instructional time (Bellei, 2009; Lavy, 2012).³ Our results are robust to accounting for potential sorting based on subject-specific student unobservables and restricting the sample to students taught by the same teacher in both subjects, thus also holding constant any teacher characteristics that do not differ across subjects.

Exploiting the cross-country dimension of our data and vast differences in economic development—with GDP per capita varying by a factor of 30—we find that teacher subject knowledge is only effective in more developed countries. As discussed in Hanushek, Link and Woessmann (2013), GDP per capita reflects resources and the quality of those institutions that promote productivity and social interaction. Similarly, we find that the impact of teacher subject knowledge is larger in well-equipped schools and for students with access to subject-specific textbooks.

Our work is related to the literature on the determinants of student achievement, which mostly deals with developed countries. This literature shows that teachers differ greatly in their ability to enhance student learning (for a review, see Jackson, Rockoff and Staiger, 2014). However, easily-observed teacher characteristics, such as education, gender, and teaching ex-

²This estimate is based on the midpoint (= 0.15 SD) of the range of estimates on how much student performance increases when teacher value-added increases by one SD in the United States (Jackson, Rockoff and Staiger, 2014). A recent study from India finds a similar estimate (Azam and Kingdon, 2015).

³See Section 7 for a discussion of how our effect size relates to the impact of teachers in other settings and of other types of educational inputs on student performance.

perience (except for the first few years), are not consistently related to teacher effectiveness (Hanushek and Rivkin, 2006). The only teacher trait consistently associated with gains in student performance is teacher cognitive skills, as measured by achievement tests (e.g., Eide, Goldhaber and Brewer, 2004; Hanushek, 1986; Hanushek and Rivkin, 2006; Rockoff et al., 2011).⁴ Hanushek, Piopiunik and Wiederhold (2016) also find positive effects of teacher cognitive skills on student achievement across OECD countries. However, in contrast to this study, the authors cannot match students to their teachers, but instead rely on country-level measures of teacher skills. Moreover, our measures of teacher subject knowledge reflect the knowledge that is essential for teaching the curriculum, and therefore differ considerably from the more general teacher ability measures employed in most of the previous literature.

In the context of developing countries, several studies find positive correlations between teacher test scores and student achievement.⁵ However, these studies likely suffer from bias due to omitted student and teacher characteristics and non-random sorting of students and teachers. Metzler and Woessmann (2012) circumvent these problems by exploiting within-teacher, within-student variation across two subjects for sixth-grade students and their teachers in Peru, finding a significant impact of teacher skills on student achievement. In contrast to our study, the authors focus on a single country and therefore cannot investigate potential effect heterogeneity by country characteristics.⁶

The remainder of the paper is structured as follows. Section 2 describes the data. Section 3 lays out the estimation strategy. Section 4 presents the results regarding the effect of teacher subject knowledge on student learning, while Section 5 explores the robustness of these findings. Section 6 presents results on effect heterogeneity. Section 7 compares the magnitude of the teacher subject knowledge impact to other settings and to other education inputs, including teacher incentives. Section 8 concludes.

⁴The evidence for teachers' scores on licensure tests affecting student performance is mixed (Clotfelter, Ladd and Vigdor, 2006; Harris and Sass, 2006; Goldhaber, 2007).

⁵Examples include Santibañez (2006) for Mexico, Behrman, Ross and Sabot (2008) for Pakistan, and Marshall (2009) for Guatemala. Behrman (2010), Glewwe et al. (2013), and Murnane and Ganimian (2014) provide overviews of the literature.

⁶Three other studies attempt to identify the impact of teacher subject knowledge on student performance using the SACMEQ data, but they are substantially different from our paper. Shepherd (2015) restricts her attention to a single country (South Africa) and Altinok (2013) uses a simple OLS model without student fixed effects. Hein and Allen (2013) focus primarily on other teacher characteristics such as experience.

2 Data and Descriptive Statistics

2.1 The SACMEQ Assessments

The empirical analysis draws on data from the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), a collaborative network of 15 Sub-Saharan African ministries of education and the UNESCO International Institute for Educational Planning (IIEP). The network periodically conducts international assessments of the math and reading knowledge of sixth-grade primary-school students and their teachers. By means of student, teacher, and principal questionnaires, it also collects detailed background information on student and teacher characteristics, as well as on classroom and school resources. The first wave of this assessment took place in 1995 and covered seven countries; the second wave, in 2000, covered 14 countries; and the third wave, in 2007, covered 15 countries. In this paper, we use data from the last two waves because teachers were not tested in the first wave.

SACMEQ employs a two-stage clustered sampling design to draw nationally representative samples of sixth-grade students for each participating country. Schools are sampled within pre-defined geographical strata in the first stage, and a simple random sample of students is drawn from each selected school in the second stage. In the second wave, 20 students per school were sampled randomly, and the teachers who taught math and reading to these students were tested. In the third wave, 25 students per school were sampled randomly, and the math and reading teachers of the three largest classes in each school were tested.⁷ While all students are tested in both math and reading, teachers are tested only in the subject they teach. However, both math and reading scores are available for a subsample of teachers who teach sampled students in both subjects.

The SACMEQ student assessments are designed to reflect the elements common to the math and language curricula in the participating countries. The multiple-choice tests contain items developed by SACMEQ itself as well as items from other international student assessments

⁷The sampling design of the third wave implies that teacher test scores are missing for students who did not attend any of the three largest classes. As explained in Section 2.2, all students with missing teacher test scores are excluded from the sample.

such as the Trends in International Mathematics and Science Study (TIMSS). Students in all participating countries are administered the same tests at the end of sixth grade, with tests translated into the local language of instruction if it is different from English. The teacher assessments include items from the student tests and additional, more difficult questions.

Both student and teacher tests are graded centrally in each country under the auspices of the IIEP. Using Item Response Theory, all test scores are placed on a common scale with a mean of 500 and a standard deviation of 100 across students participating in the second SACMEQ wave. Because of the overlapping items, test scores are directly comparable between students and teachers as well as between the two assessment waves. The similarity between student and teacher tests also means that teacher test scores in SACMEQ reflect knowledge that is likely highly relevant for teaching math and reading. Therefore, these curriculum-based measures of teacher knowledge are noticeably different from other teacher test scores, for instance, SAT and ACT scores in the United States, which reflect teachers' general cognitive ability.

2.2 Sample Selection and Descriptive Statistics

We pool the data from the second and third waves of the SACMEQ assessment. We exclude Mauritius from the original set of 15 countries because it did not test teachers. Teachers in South Africa were not tested in the second wave and could opt out in the third wave, which 18% of the sampled teachers did. Thus, we also exclude South Africa from the analysis.⁸ We further exclude 5,428 students who could not be linked to a teacher in any subject, 4,018 students who had at least one teacher with missing test scores, and 225 students with missing test scores.⁹ The final sample consists of 74,708 students with 8,742 teachers in 3,939 schools in 13 countries: Botswana, Kenya, Lesotho, Malawi, Mozambique, Namibia, Seychelles, Swaziland, Tanzania

⁸Opting out (by either students or teachers) was not possible in any other country. Results are robust to retaining South Africa in the sample.

⁹Some background variables have missing values. Since we consider a large set of explanatory variables and since a portion of these variables is missing for a relatively large fraction of students, dropping all student observations with any missing value would result in substantial sample reduction. We therefore imputed missing values for control variables by using the country-by-wave means. To ensure that imputed data are not driving our results, all our regressions include an indicator for each variable with missing data that equals 1 for imputed values and 0 otherwise.

(mainland), Uganda, Zambia, Zanzibar (semi-autonomous region of Tanzania), and Zimbabwe (participated only in the third wave).

Table A-1 reports descriptive statistics of student performance and teacher subject knowledge for the pooled sample and separately for each country. There are striking differences in student performance between countries. For example, in math, students in Kenya score on average more than 1.4 international SD higher than students in Zambia. Similarly, in reading, students in the Seychelles score more than 1.5 international SD higher than their peers in Malawi. Interestingly, the cross-country differences in teacher subject knowledge are even larger. Teachers in Kenya, for example, outperform teachers in Zanzibar by 2.2 international SD in math; the variation in teacher reading knowledge is of a similar magnitude.¹⁰ Figures A-1 and A-2 illustrate these large cross-country differences by plotting each country's distribution of teacher test scores in math and reading and, as a benchmark, the average test score of teachers in the best-performing country.

To put the observed variation in teacher subject knowledge into perspective, we compare it to the subject-knowledge variation between teachers with different levels of education. For instance, in the pooled sample, the average math test score is 734 points for teachers with only primary education and 822 points for teachers with tertiary education. This difference is equivalent to 0.8 international SD in teacher subject knowledge in math. In other words, the difference in teacher math knowledge between the country with the best-performing teachers and the country with the worst-performing teachers is almost three times as large as the difference between teachers with tertiary education and teachers with primary education (in reading, this ratio is about 2). Another way to illustrate the substantial differences in teacher subject knowledge across countries is to consider individual test items. For instance, teachers participating in SACMEQ were asked to answer the following math question: “ $x/2 < 7$ is equivalent to (a) $x > 14$, (b) $x < 14$, (c) $x > 5$, or (d) $x < 7/2$?” Eighty-three percent of teachers in Kenya answered this question correctly, but only 43% of teachers in Lesotho did so.¹¹

¹⁰In each country, the average teacher significantly outperforms the average student in both math and reading. However, in all countries, the best students outperform the worst teachers.

¹¹There are even bigger cross-country differences for the following item: “If the height of a fence is raised from 60cm to 75cm, what is the percentage increase in height: (a) 15%, (b) 20%, (c) 25%, or (d) 30%?” Correct answer rates vary between 18% in Zanzibar and 88% in Kenya.

These large cross-country differences notwithstanding, teachers in Sub-Saharan Africa have much less knowledge than teachers in developed countries. While there is no dataset that would allow a direct comparison between African teachers and teachers in developed countries, we can compare the math knowledge of teachers in Sub-Saharan Africa to that of eighth-grade students in developed countries. In the TIMSS 1995 assessment, eighth-grade students were asked to solve the same math question described above (“ $x/2 < 7$ is equivalent to”). In 19 out of 39 mostly developed countries, eighth-grade students did as well or even better than teachers in the worst-performing Sub-Saharan country (Lesotho), and in four countries they did even better than the average teacher in Sub-Saharan Africa. Moreover, 47% of eighth-grade students in the United States could solve this math problem, and—judging by this item alone—are therefore at the level of teachers in Botswana and Namibia.¹²

2.3 Relationship Between Student Performance and Teacher Subject Knowledge at the Country Level

To get a first sense of the importance of teacher subject knowledge for student performance, we plot average student test scores against average teacher test scores at the country level. The upper panel of Figure 1 reveal positive associations for both math and reading: students in countries with highly knowledgeable teachers tend to perform better than their peers in countries with teachers who have less of a command of the material they are teaching.

The availability of both student and teacher performance measures is a unique feature of the SACMEQ assessments. Other international student assessments contain at best coarse measures of teacher quality, for example, teachers’ educational attainment. To understand if teacher subject knowledge is a better predictor of student performance than their credentials, the bottom panel of Figure 1 plots a country’s average student performance against the share

¹²These comparisons actually overestimate the relative performance of teachers in Sub-Saharan Africa because, in the SACMEQ assessment, the teachers had only four different answers from which to choose, whereas the eighth-grade students in TIMSS had to choose among five possible answers.

of teachers with a college degree. Unlike subject knowledge, educational credentials appear to explain little if any of the cross-country variation in student performance.¹³

3 Estimation Strategy

In the baseline OLS model, we estimate the following education production function:

$$y_{ikcs} = \alpha + \beta T_{kcs} + \gamma_1 X_{ics} + \gamma_2 X_{cs} + \gamma_3 X_s + \delta Z_{kcs} + \xi_c + \varepsilon_{ikcs}, \quad (1)$$

where y_{ikcs} is the test score of student i in subject k (math or reading) in classroom c in school s ; T_{kcs} is the test score of student i 's teacher in subject k ; X_{ics} is a vector of student-level subject-invariant controls measuring student and family background; X_{cs} is a vector of subject-invariant classroom and teacher characteristics; X_s is a vector of subject-invariant school characteristics; and Z_{kcs} contains classroom and teacher characteristics that vary across subjects (e.g., the availability of teacher guides in math or reading).¹⁴ ξ_c is a vector of country fixed effects, which absorb any country-specific differences in student performance.¹⁵ ε_{ikcs} is the error term.¹⁶

Interpreting the OLS estimate of β as the causal effect of teacher subject knowledge on student performance is problematic because of omitted variables that might be correlated with both student and teacher test scores. For instance, $\hat{\beta}$ would be biased upward if highly educated parents select schools or classrooms with better teachers and also foster their children's learning in other ways. Similarly, student sorting across or within schools would lead to biased estimates if students with high (unobserved) academic ability are more likely to attend schools or classrooms with highly knowledgeable teachers.

To address these sources of bias, we exploit the fact that students were tested in two subjects and ask whether *differences* in teacher knowledge between math and reading are systematically related to *differences* in student performance between the same two subjects. This implies that

¹³A qualitatively similar picture emerges if we instead use the share of teachers who completed at least secondary school.

¹⁴See Table 1 for a complete list of control variables.

¹⁵The country fixed effects also control for potential cross-country differences in school curricula or in the timing of national examinations.

¹⁶Additionally, we include a wave dummy in all specifications. To simplify notation, we omit the wave dummy and the wave subscripts in all equations.

we identify the effect of teacher subject knowledge based only on variation between teacher math and reading knowledge within the same student.¹⁷ We thus estimate the following first-differenced model, which we implement by pooling the two subjects math and reading and adding student fixed effects to Equation (1):

$$\begin{aligned} y_{ics,math} - y_{ics,read} &= \beta(T_{cs,math} - T_{cs,reading}) \\ &\quad + \delta(Z_{cs,math} - Z_{cs,reading}) + (\varepsilon_{ics,math} - \varepsilon_{ics,reading}). \end{aligned} \tag{2}$$

This model controls for the influence of any student-level performance determinants that are not subject-specific, such as family background, overall academic ability, or general motivation. It also eliminates the impact of school resources that do not differ across subjects, such as availability of black boards, chairs, and computers. Therefore, estimates from the student fixed effects model are not biased by student sorting across or within schools, as long as such sorting is not subject-specific.¹⁸ In robustness checks, we provide evidence that our estimates are also unlikely to be biased by subject-specific sorting.

The within-student model of Equation (2) ensures that the estimates are not confounded by any subject-invariant student characteristics; however, unobserved teacher traits could still bias the coefficient on teacher subject knowledge. For example, if teachers with high subject knowledge are also more motivated (not observed in the data), a positive estimate of β might partly reflect the impact of high motivation. The fact that about one-third of the students in our sample are taught both math and reading by the same teacher allows us to address this issue in a robustness check. Specifically, by restricting the sample to students taught both subjects by the same teacher (*same-teacher sample*), we can control for any teacher traits that affect students' math and reading performance in the same way.¹⁹ The results suggest that our student fixed effects estimates are not confounded by correlated teacher traits.

¹⁷Within-student across-subject variation has been exploited in previous studies (e.g., Dee, 2005, 2007; Clotfelter, Ladd and Vigdor, 2010; Lavy, 2015).

¹⁸In contrast to the OLS model, the impact of teacher subject knowledge in the fixed effects model is “net” of teacher knowledge spillovers across subjects.

¹⁹Using the same-teacher sample is equivalent to adding teacher fixed effects in Equation (2), thus exploiting only variation within students and within teachers.

4 Results

4.1 Ordinary Least Squares Results

Table 1 reports estimates of the association between student performance and teacher subject knowledge in math (Panel A) and in reading (Panel B) based on the model in Equation (1). In addition to an increasing number of control variables at the student, classroom, school, and teacher level, all specifications include country fixed effects.²⁰ To facilitate interpretation of effect sizes, both student and teacher test scores are standardized with a mean of 0 and a standard deviation of 1 across countries and waves. Throughout our analysis, we cluster standard errors at the school level.²¹

The results in Table 1 show a strong positive association between teacher subject knowledge and student performance in both math and reading. In the most parsimonious specification that includes only country fixed effects, a 1 SD increase in teacher subject knowledge is associated with a 0.12 SD increase in student performance in both subjects (Column 1). This association becomes weaker when student, classroom, and school characteristics are added as controls, but remains statistically significant (Columns 2–4). Interestingly, the coefficient on teacher subject knowledge changes only slightly when teacher characteristics, such as educational attainment and experience, are also controlled for (Column 5). In this most restrictive specification, a 1 SD increase in teacher subject knowledge is associated with a 0.07 (0.06) SD increase in student performance in math (reading).²²

²⁰Because the regressions in Table 1 use only within-country variation, the coefficients do not correspond to the cross-country correlations in the upper panel of Figure 1.

²¹The SACMEQ data include student sampling weights, and we confirmed that our coefficient estimates are virtually identical independently of whether we weight observations or not in the regressions. However, as described in Solon, Haider and Wooldridge (2015), using sampling weights may unnecessarily decrease precision, an issue that affects especially our regressions that focus on smaller sub-samples of students. We therefore chose to report the unweighted estimates in the paper.

²²An assumption embodied in the student fixed effects model is that the effect of teacher subject knowledge is similar across subjects. Supporting this assumption, a cross-equation test indicates that one cannot reject the equality of OLS coefficients in math and reading (in the full-control models in Column 5, the respective p-value is 0.211).

4.2 Student Fixed Effects Results

The OLS estimates in Table 1 are likely biased due to omitted variables and non-random sorting across or within schools. Therefore, we now turn to the student fixed effects model that identifies the impact of teacher subject knowledge based only on within-student variation between math and reading. The results, shown in Table 2, indicate that teacher subject knowledge has a positive and statistically significant impact on student performance. When controls for subject-specific classroom and teacher characteristics are added, a 1 SD increase in teacher subject knowledge raises student performance by 0.026 SD (Column 3).²³ This suggests that differences in teacher subject knowledge account for about 20% of the variation in teacher value added, with evidence on teacher value added coming from India and the United States (Chetty, Friedman and Rockoff, 2014; Jackson, Rockoff and Staiger, 2014; Azam and Kingdon, 2015).

Compared to the OLS estimates in Table 1, the fixed effects estimate on teacher subject knowledge is much smaller. One obvious explanation for this finding is that unobserved student characteristics correlated with both student and teacher test scores bias the OLS estimates upward. Another possible explanation is that attenuation bias due to measurement error in teacher subject knowledge is aggravated in the fixed effects model (see Angrist and Krueger, 1999, Chapter 4). In the Appendix, we show how the reliability ratios of the teacher math and reading tests can be used to correct for measurement error. We find that the measurement-error-corrected coefficient on teacher subject knowledge is 50% larger than the baseline estimate (0.039 SD vs. 0.026 SD). However, this correction procedure hinges on several strong assumptions, such as that the measurement errors in math and reading tests are uncorrelated. Therefore, we report only the uncorrected, more conservative, estimates throughout the paper.²⁴

Figure 2 shows a non-parametric version of the regression in Column 3 of Table 2. To create this binned scatterplot, we first regressed the differences between math and reading scores of

²³Besides teacher subject knowledge, the only statistically significant explanatory variables are a dummy for female teachers and a dummy for teachers having access to a teaching guide for their subject; the coefficients on both variables are positive.

²⁴Note that differential measurement error in teacher math and reading knowledge could create bias in the student fixed effects model. Reassuringly, however, test reliability in math and reading is very similar, with estimated reliability ratios of 0.83 for math and 0.75 for reading (see Appendix), thus indicating a similar degree of measurement error in both subjects.

students and the respective test score differential of teachers separately on all control variables (also in differences). We then divided the residualized teacher test scores into 20 equal-sized groups and plotted the mean value in each bin against the mean value of the residualized student test scores. The figure suggests that the relationship between the test score differentials is roughly linear. To illustrate this result, consider two teachers: one teacher has average knowledge in both subjects (e.g., math=0, reading=0), and the other teacher has higher math than reading knowledge (e.g., math=1, reading=0). Suppose that the math knowledge of both teachers improves by 1 SD. Then, the relative math performance (vs. reading performance) of the students of both teachers increases by the same amount.

One important question concerning the interpretation of our results is whether the estimates capture the impact of teacher subject knowledge for only a single school year or, instead, the cumulative effect over several school years. Unfortunately, the SACMEQ data do not contain information on how long each teacher has been teaching a particular class. However, there is ample evidence that teacher turnover in Sub-Saharan Africa is high, with annual attrition rates ranging between 5-30% (Mulkeen et al., 2007). Moreover, a study from two Malawian school districts finds that almost 50% of the 188 teachers who began the school year were not teaching the same class nine months later (IEQ, 2000). Given this high turnover in the teacher workforce, our estimates likely capture a short-run effect of teacher subject knowledge on student performance.

5 Robustness Checks

5.1 Subject-Specific Student Sorting

The student fixed effects specifications identify the impact of teacher subject knowledge based only on within-student between-subject variation. Thus, they account for potential sorting of students to schools or teachers based on subject-invariant factors, such as students' overall academic ability. The estimates will be biased, however, if sorting is based on subject-specific factors. For example, our estimate will be biased upward if mathematically gifted students systematically attend schools with knowledgeable math teachers or if principals tend to assign

mathematically gifted students to teachers with high math knowledge. Columns 1 to 4 of Table 3 suggest that these mechanisms are unlikely to drive our results.

We first address the issue of sorting *across* schools by restricting the sample to students living in rural areas, where students likely have little choice between different schools. Column 1 of Table 3 shows that the estimated coefficient on teacher subject knowledge in this sample is similar to our baseline coefficient, suggesting that it is unlikely that non-random sorting of students across schools is biasing our results. To address the concern of sorting *within* schools, we focus on schools with only one sixth-grade classroom, meaning that principals cannot assign students to teachers based on their subject-specific ability. As shown in Column 2, the impact of teacher subject knowledge in this sample is similar to the estimate in the full sample. Column 3 shows that our results hold even when we restrict the sample to one-classroom schools in rural areas, simultaneously addressing across-school and within-school sorting.

An alternative way of accounting for potential sorting based on subject-specific factors within schools is to aggregate teachers' subject knowledge to the school level. Again, our estimate remains unaffected, suggesting that non-random sorting within schools is not driving our results (Column 4 of Table 3).²⁵ Finally, a particularly salient motive for subject-specific sorting relates to the match between the language spoken by students at home and at school.²⁶ For example, students who speak English at home ("native speakers") may have a preference for schools with teachers who are proficient in English. However, we investigated whether the impact of teacher subject knowledge varies with the share of native speakers at school and found no evidence of such effect heterogeneity (for details, see Bietenbeck, Piopiunik and Wiederhold, 2015).

5.2 Unobserved Teacher Traits

Another concern is that our estimates reflect the effect of other, unobserved teacher characteristics correlated with subject knowledge. For example, teachers with high subject knowledge

²⁵For this analysis, all other teacher characteristics are aggregated to the school level as well.

²⁶In the countries covered by SACMEQ, English is typically both the official language of instruction and the test language. In practice, however, the dominant language of instruction and the language which students use in their daily lives may not be English.

might also have excellent pedagogical skills, which would bias the coefficient of interest upward. In Column 5 of Table 3, we address this concern by restricting the sample to students taught both math and reading by the same teacher (which is equivalent to including teacher fixed effects in the full sample).²⁷ Identification in this same-teacher sample is based only on variation in subject knowledge between math and reading *within* teachers. Hence, all subject-invariant teacher traits, such as general pedagogical skills or absenteeism, are controlled for. Using the same-teacher sample leaves our baseline student fixed effects results unchanged, indicating that unobserved subject-invariant teacher characteristics are unlikely to bias our estimates.²⁸

5.3 Subject-Specific Impacts of Covariates

One of the assumptions underlying the student fixed effects model is that subject-invariant covariates (e.g., parental education) have a similar impact on student performance in both math and reading. To account for the possibility that covariates affect math and reading performance differently, we estimate a model in which all student and school characteristics are interacted with a subject dummy.²⁹ As Column 6 of Table 3 shows, the estimated coefficient on teacher subject knowledge is slightly smaller but remains statistically significant.³⁰

²⁷ Across all countries in our sample, 35% of students are taught both subjects by the same teacher.

²⁸ While we control for any differences between teachers that are similar across subjects—most importantly, general motivation and pedagogical skills—our results might still be affected by subject-specific teacher traits (e.g., particularly high motivation in one subject) if correlated with subject knowledge. We do not have information on subject-specific teacher effort or motivation, and hence, we cannot include such measures in the set of controls. However, it seems likely that differences in unobserved teacher traits are much larger between teachers than within the same teacher across the two subjects. While we cannot test this assumption directly, our data allow us to assess between-teacher vs. within-teacher variation in *observable* teacher traits, that is, teacher subject knowledge. Using the math and reading scores of teachers observed in both subjects (i.e., our same-teacher sample), we find that 71 percent of the test score variation is *between* teachers and only 29 percent is *within* teachers. Still, just as in Metzler and Woessmann (2012), our results should be interpreted as the impact of teacher subject knowledge and any subject-specific trait correlated with it.

²⁹ Results are similar when additionally including interactions of the ten family resources with the subject dummy.

³⁰ While we would also like to control for subject-specific instructional time as a potentially important determinant of student performance, our data do not provide this information. Existing evidence suggests that this is not a major concern since Metzler and Woessmann (2012) show for Peru that including subject-specific instruction time does not affect the estimate of the effect of teacher subject knowledge on student performance.

In Column 7, we additionally allow country-specific impacts to differ across math and reading by including subject-by-country fixed effects. In this specification, identification is based only on within-student variation in teacher subject knowledge relative to the country mean. Since the between-country variation in both teacher subject knowledge and student performance is substantial (see Table A-1), this implies that a considerably smaller part of the sample variation is used. The impact of teacher subject knowledge remains positive in this model, although the coefficient is substantially smaller than our main estimate. This suggests that the impact of teacher subject knowledge on student performance varies substantially across countries, an issue that we investigate in the next section.

6 Heterogeneity

A unique feature of the SACMEQ data is that they provide comparable measures of teacher subject knowledge and student performance for a relatively large number of countries. Furthermore, the countries included in SACMEQ differ substantially in regard to economic development. For example, GDP per capita ranges from \$595 in Mozambique to \$17,811 in the Seychelles, a difference by a factor of 30.³¹ This allows us to investigate whether the impact of teacher subject knowledge varies systematically with a country's level of economic development.

In Table 4, we find that teacher subject knowledge is effective only in more developed countries, as measured by higher GDP per capita (Column 1) or a higher rank on the Human Development Index (HDI), which is a broader measure based on income, life expectancy, and literacy (Column 2).³² In countries at a higher stage of development, a 1 SD increase in teacher subject knowledge increases student learning by 0.04–0.05 SD. While a causal interpretation of these results is clearly difficult as countries differ along several unobserved dimensions,

³¹Figures are averages of the SACMEQ assessment years 2000 and 2007 and are measured in PPP-US-dollars.

³²The similarity of results for GDP per capita and HDI rank is not surprising since both indicators are highly correlated in our country sample ($r=0.92$).

the remainder of Table 4 digs deeper into one potential mechanism explaining this country heterogeneity.³³

One reason why the effect of teacher subject knowledge varies with the country's stage of development may be that richer countries have more resources to spend on schools.³⁴ SACMEQ provides various measures of school resources at the student and school level, allowing us to test more directly whether the estimated teacher effect varies with a school's resource endowment. We first interact teacher subject knowledge with textbook availability during class, a crucial educational resource that is often lacking in Sub-Saharan Africa. Since each student reports the availability of textbooks separately for math and reading, we exploit within-student across-subject variation in both teacher knowledge and textbook availability.³⁵ Column 3 of Table 4 shows that an increase in teacher subject knowledge improves student performance twice as much for students who have textbooks during class compared to students without textbooks.

In Columns 4 and 5 of Table 4, we consider more general school resources. SACMEQ contains information on the availability of a large variety of school resources (reported by principals), ranging from blackboards, chairs, and tables to access to drinking water. We combine all 31 school resources into a single index by counting the number of available resources.³⁶ Column 4 suggests that teacher knowledge is more effective for student learning when more school resources are available. In contrast, we find no significant interaction between teacher

³³In additional analysis, we also estimated the impact of teacher subject knowledge separately for each country. We find positive point estimates in almost all countries. However, due to the small sample sizes, we could detect a significant effect in only 2 of 13 countries.

³⁴Another reason might be that GDP per capita reflects the quality of (educational) institutions (Hanushek, Link and Woessmann, 2013). Other reasons could include differences between countries involving teacher absenteeism, teacher effort, and/or learning culture.

³⁵Students were asked "How are the *math* textbooks used in your classroom during the lessons?", with five answer categories: (1) There are no math textbooks; (2) Only the teacher has a math textbook; (3) I share a math textbook with two or more pupils; (4) I share a math textbook with one pupil; (5) I use a math textbook by myself. The analogous question was asked about *reading* textbooks. In line with Glewwe, Kremer and Moulin (2009), we group students who use a textbook by themselves and students who share a textbook with only one other student because all these students can effectively use a textbook during class. This categorization is also consistent with experimental evidence from the Philippines that providing one textbook for every two students and providing one textbook for each student has very similar effects on test scores (Heyneman, Jamison and Montenegro, 1984). The sample mean of our binary textbook variable is 0.56 for math and 0.58 for reading.

³⁶To facilitate interpretation of results, we normalize all school-level variables to have mean of 0 and SD of 1, such that the main effect of teacher subject knowledge reflects the impact at the sample mean of the respective resource variable. Because school-level resources do not vary across subjects, their main effects on student performance cannot be estimated.

subject knowledge and class size, suggesting that teachers with the same level of subject knowledge are as effective in large classrooms as in small ones (Column 5). Taken together, the results in Columns 3–5 suggest a potential mechanism explaining the cross-country differences in the teacher effect: more developed countries have better school resources, which may be complementary to teacher subject knowledge in educational production.³⁷

7 Discussion of Teacher Subject Knowledge Impact

To put our results into perspective, we compare our teacher knowledge impact to effect sizes in other settings as well as to other types of education inputs and teacher incentives in (mostly) developing countries. Based on sixth-grade students in Peru and using the same identification strategy as is employed in this paper, Metzler and Woessmann (2012) find that a 1 SD increase in teacher knowledge raises student achievement by about 0.04 SD. This is similar to our estimated impact of teacher subject knowledge (0.03 SD). Teachers' overall impact on student achievement is commonly estimated in value-added (VA) models. Based on administrative data from the United States, Chetty, Friedman and Rockoff (2014) estimate that a 1 SD improvement in teacher VA raises student achievement by about 0.14 SD in math and 0.1 SD in English. Assuming that the variation in overall teacher effectiveness in Sub-Saharan Africa is similar to that in the United States, this implies that teacher subject knowledge explains between 20% (in math) and 25% (in reading) of the variation in teachers' overall effectiveness.

Several previous studies evaluate interventions that aimed to increase student achievement by providing schools with teaching inputs, such as textbooks (Glewwe, Kremer and Moulin, 2009; Sabarwal, Evans and Marshak, 2014) and flipcharts (Glewwe et al., 2004). These inputs typically fail to improve student achievement, either because they are not used or because teachers cannot use them effectively. Hence, the impact of teacher subject knowledge is larger than providing these resources.

Given that many students in low-income countries attend school for only half the day, another approach to improve children's performance is to expand instructional time. In evaluating

³⁷In line with this interpretation, we observe that GDP per capita is strongly correlated with textbook availability ($r=0.61$) and the index of school resources ($r=0.94$) at the country level.

the full school day program in Chile, Bellei (2009) finds that a reform that, among other things, increased the number of instruction hours per day by 10% improved the Spanish and math achievement of 10th graders by 0.02 SD and 0.03 SD, respectively. Lavy (2012) similarly finds that a 10% expansion in the weekly hours devoted to English, math, and science in Israel improved the achievement of fifth graders in these subjects by 0.03 SD. Increasing time spent on instruction thus has an impact on student achievement very similar to that achieved by improving teacher subject knowledge by 1 SD.

The extant literature also studies how incentives that reward teachers for additional effort or for improving their students' test scores affect student performance. Duflo, Hanna and Ryan (2012) evaluated a program in India that rigorously monitored teacher attendance and offered them bonuses based on the number of days they attended school. This program strongly reduced teacher absenteeism and improved student achievement in math and Hindi by 0.17 SD after 30 months. Muralidharan and Sundararaman (2011) find that offering teachers cash incentives for improving their students' performance in standardized tests raised test scores in math and Telugu by 0.14 and 0.16 SD, respectively, after one year. Unconditional pay increases for teachers, in contrast, do not seem to work (de Ree et al., 2015). Compared to well-designed (and costly) teacher incentive schemes, the impact of teacher subject knowledge is therefore rather small.

The above comparisons reveal that the impact of teacher subject knowledge on student performance falls somewhere between that of interventions with zero impact (e.g., providing textbooks or flipcharts) and interventions with strong impacts (e.g., paying teachers for attendance that is rigorously monitored).

8 Conclusion

Student performance in Sub-Saharan Africa is low, which may partly explain the region's poor economic performance, given that the cognitive skills of a population are an important driver of economic growth (Hanushek and Woessmann, 2012). We investigate the role of teacher quality in explaining the low student performance, focusing on teacher subject knowledge as

one central dimension of teacher quality. Our measures for teacher knowledge in math and reading are curriculum-based, thus reflecting the subject knowledge required for teaching. To identify a causal effect of teacher subject knowledge, we exploit within-student variation across math and reading. We find that a 1 SD increase in teacher subject knowledge raises student performance by 0.03 SD. Results are robust to including teacher fixed effects and are not driven by sorting of students or teachers. Exploiting the vast differences in the countries' economic development, we also provide suggestive evidence that teacher subject knowledge is effective in improving student learning only in countries at a higher stage of development.

Although the effects of increasing teacher subject knowledge on student performance are modest, they are comparable to other well-known interventions, such as expanding instructional time. Moreover, the low skills of teachers in developing countries may limit the impact of other educational interventions (e.g., when it comes to using textbooks effectively). Hence, it seems essential to increase the skills of the teacher workforce, either by improving the skills of existing teachers or by hiring teachers with better skills.

References

- Altinok, Nadir.** 2013. “The Impact of Teacher Knowledge on Student Achievement in 14 Sub-Saharan African Countries.” *Mimeo*.
- Angrist, Joshua D., and Alan B. Krueger.** 1999. “Empirical Strategies in Labor Economics.” In *Handbook of Labor Economics*. Vol. 3 Part A, , ed. David Card and Orley Ashenfelter, 1277–1366. Elsevier.
- Azam, Mehtabul, and Geeta Kingdon.** 2015. “Assessing Teacher Quality in India.” *Journal of Development Economics* 117: 74–83.
- Behrman, Jere R.** 2010. “Investment in Education: Inputs and Incentives.” In *Handbook of Development Economics*. Vol. 5, , ed. Dani Rodrik and Mark Rosenzweig, 4883–4975. Elsevier.
- Behrman, Jere R., David Ross, and Richard Sabot.** 2008. “Improving Quality Versus Increasing the Quantity of Schooling: Estimates of Rates of Return from Rural Pakistan.” *Journal of Development Economics* 85: 94–104.
- Bellei, Cristián.** 2009. “Does lengthening the school day increase students’ academic achievement? Results from a natural experiment in Chile.” *Economics of Education Review* 28: 629–640.
- Bietenbeck, Jan, Marc Piopiunik, and Simon Wiederhold.** 2015. “Africa’s Skill Tragedy: Does Teachers’ Lack of Knowledge Lead to Low Student Performance?” CESifo Working Paper No. 5470.
- Chetty, Raj, John N. Friedman, and Jonah E. Rockoff.** 2014. “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates.” *American Economic Review* 104(9): 2593–2632.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. “Teacher-Student Matching and the Assessment of Teacher Effectiveness.” *Journal of Human Resources* 41: 778–820.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2010. “Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects.” *Journal of Human Resources* 45: 655–681.

Dee, Thomas S. 2005. “A Teacher like Me: Does Race, Ethnicity, or Gender Matter?” *American Economic Review, Papers and Proceedings* 95: 158–165.

Dee, Thomas S. 2007. “Teachers and the Gender Gaps in Student Achievement.” *Journal of Human Resources* 42: 528–554.

de Ree, Joppe, Karthik Muralidharan, Menno Pradhan, and Halsey Rogers. 2015. “Double for Nothing? Experimental Evidence on the Impact of an Unconditional Teacher Salary Increase on Student Performance in Indonesia.” NBER Working Paper 21806.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. “Incentives Work: Getting Teachers to Come to School.” *American Economic Review* 102: 1241–1278.

Eide, Eric, Dan Goldhaber, and Dominic Brewer. 2004. “The Teacher Labour Market and Teacher Quality.” *Oxford Review of Economic Policy* 20: 230–244.

Glewwe, Paul, Eric A. Hanushek, Sarah Humpage, and Renato Ravina. 2013. “School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010.” In *Education Policy in Developing Countries.* , ed. Paul Glewwe, 13–64. University of Chicago Press.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. “Many Children Left Behind? Textbooks and Test Scores in Kenya.” *American Economic Journal: Applied Economics* 1: 112–135.

- Glewwe, Paul, Michael Kremer, Sylvie Moulin, and Eric Zitzewitz.** 2004. “Retrospective vs. prospective analyses of school inputs: The case of flip charts in Kenya.” *Journal of Development Economics* 74: 251–268.
- Goldhaber, Dan.** 2007. “Everyone’s Doing It, But What Does Teacher Testing Tell Us About Teacher Effectiveness?” *Journal of Human Resources* 42: 765–794.
- Hanushek, Eric A.** 1986. “The Economics of Schooling: Production and Efficiency in Public Schools.” *Journal of Economic Literature* 24: 1141–1177.
- Hanushek, Eric A., and Ludger Woessmann.** 2012. “Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation.” *Journal of Economic Growth* 17: 267–321.
- Hanushek, Eric A., and Steven G. Rivkin.** 2006. “Teacher Quality.” In *Handbook of the Economics of Education*. Vol. 2, , ed. Eric A. Hanushek and Finis Welch, 1051–1078. Elsevier.
- Hanushek, Eric A., Marc Piopiunik, and Simon Wiederhold.** 2016. “The Value of Smarter Teachers: International Evidence on Teacher Cognitive Skills and Student Performance.” NBER Working Paper 20727.
- Hanushek, Eric A., Susanne Link, and Ludger Woessmann.** 2013. “Does school autonomy make sense everywhere? Panel estimates from PISA.” *Journal of Development Economics* 104(C): 212–232.
- Harris, Douglas, and Tim R. Sass.** 2006. “Value-Added Models and the Measurement of Teacher Quality.” *Mimeo*.
- Hein, Christopher F., and Rebecca Allen.** 2013. “Teacher Quality in Sub-Saharan Africa: Pupil-fixed Effects Estimates for Twelve Countries.” DoQSS Working Papers 13-08. Department of Quantitative Social Science, University College London.

- Heyneman, Stephen P., Dean T. Jamison, and Xenia Montenegro.** 1984. "Textbooks in the Philippines: Evaluation of the Pedagogical Impact of Nationwide Investment." *Educational Evaluation and Policy Analysis* 6: 139–150.
- IEQ.** 2000. "What Happened to the Teachers? Follow-Up on Teacher Mobility in Two Malawi Districts." Improving Educational Quality undertaken by: American Institutes for Research in collaboration with Academy for Educational Development, Education Development Center, Inc., Juárez and Associates, Inc., and University of Pittsburgh. Washington, D.C.
- Jackson, C. Kirabo, Jonah E. Rockoff, and Douglas O. Staiger.** 2014. "Teacher Effects and Teacher-Related Policies." *Annual Review of Economics* 6: 801–825.
- Lavy, Victor.** 2012. "Expanding School Resources and Increasing Time on Task: Effects of a Policy Experiment in Israel on Student Academic Achievement and Behavior." NBER Working Paper 18369.
- Lavy, Victor.** 2015. "Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries." *Economic Journal* 125: F397–F424.
- Marshall, Jeffery H.** 2009. "School Quality and Learning Gains in Rural Guatemala." *Economics of Education Review* 28: 207–216.
- Metzler, Johannes, and Ludger Woessmann.** 2010. "The Impact of Teacher Subject Knowledge on Student Achievement: Evidence from Within-Teacher Within-Student Variation." IZA Discussion Paper 4999.
- Metzler, Johannes, and Ludger Woessmann.** 2012. "The Impact of Teacher Subject Knowledge on Student Achievement: Evidence from Within-Teacher Within-Student Variation." *Journal of Development Economics* 99: 486–496.
- Mulkeen, Aidan, David W. Chapman, Joan G. DeJaeghere, and Elizabeth Leu.** 2007. *Recruiting, Retaining, and Retraining Secondary School Teachers and Principals in Sub-Saharan Africa.* Washington DC: World Bank.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. “Teacher performance pay: Experimental evidence from India.” *Journal of Political Economy* 119: 39–77.

Murnane, Richard J., and Alejandro J. Ganimian. 2014. “Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Evaluations.” NBER Working Paper 20284.

Rockoff, Jonah E., Brian A. Jacob, Thomas J. Kane, and Douglas O. Staiger. 2011. “Can You Recognize an Effective Teacher When You Recruit One?” *Education Finance and Policy* 6: 43–74.

Sabarwal, Shwetlena, David Evans, and Anastasia Marshak. 2014. “The Permanent Input Hypothesis: The Case of Textbooks and (No) Student Learning in Sierra Leone.” Policy Research Working Paper 7021, Washington, DC: World Bank Group.

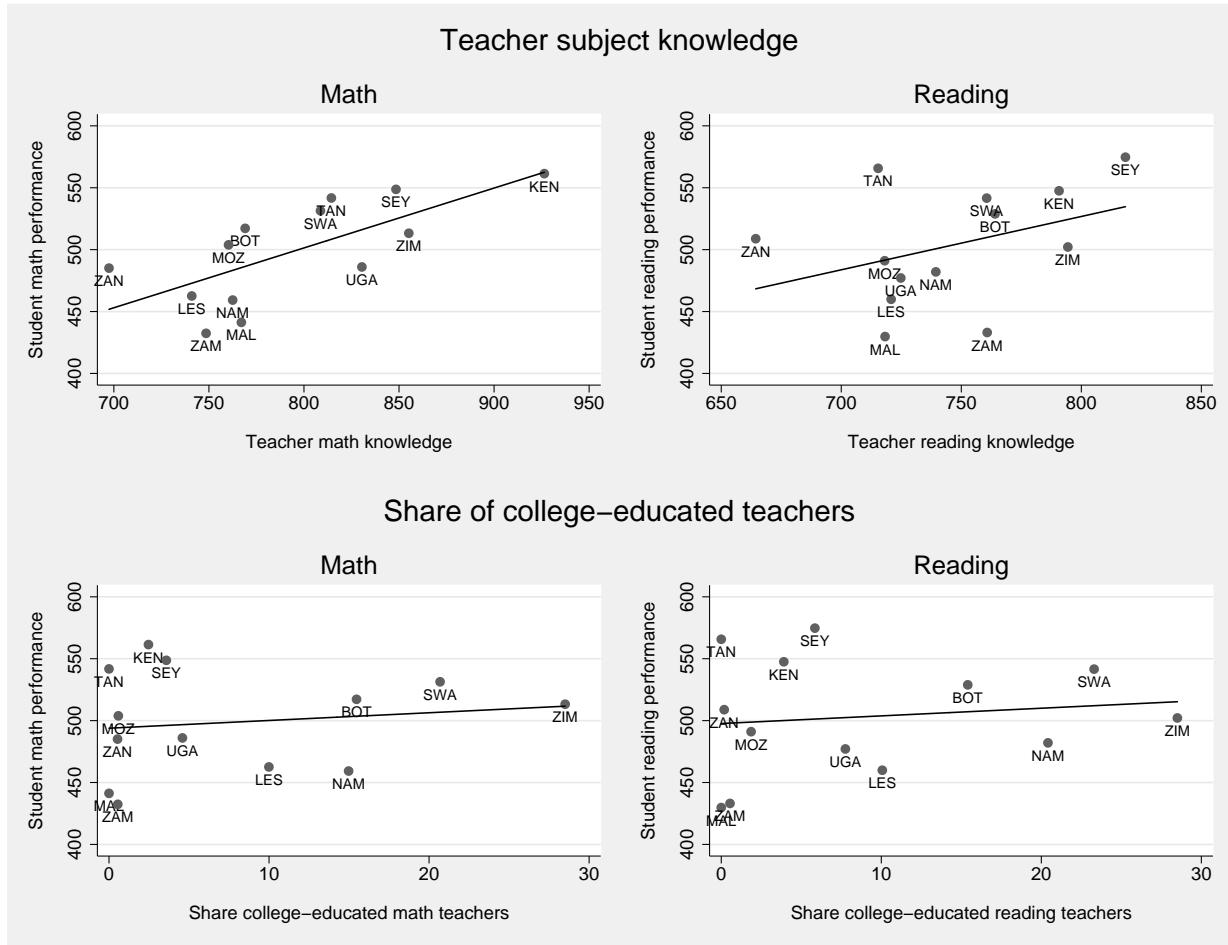
Santibañez, Lucrecia. 2006. “Why We Should Care if Teachers Get A’s: Teacher Test Scores and Student Achievement in Mexico.” *Economics of Education Review* 25: 510–520.

Shepherd, Deborah L. 2015. “Learn to Teach, Teach to Learn: A Within-Pupil Across-Subject Approach to Estimating the Impact of Teacher Subject Knowledge on South African Grade 6 Performance.” Stellenbosch Economic Working Papers 01/15.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. “What Are We Weighting For?” *Journal of Human Resources* 50(2): 301–316.

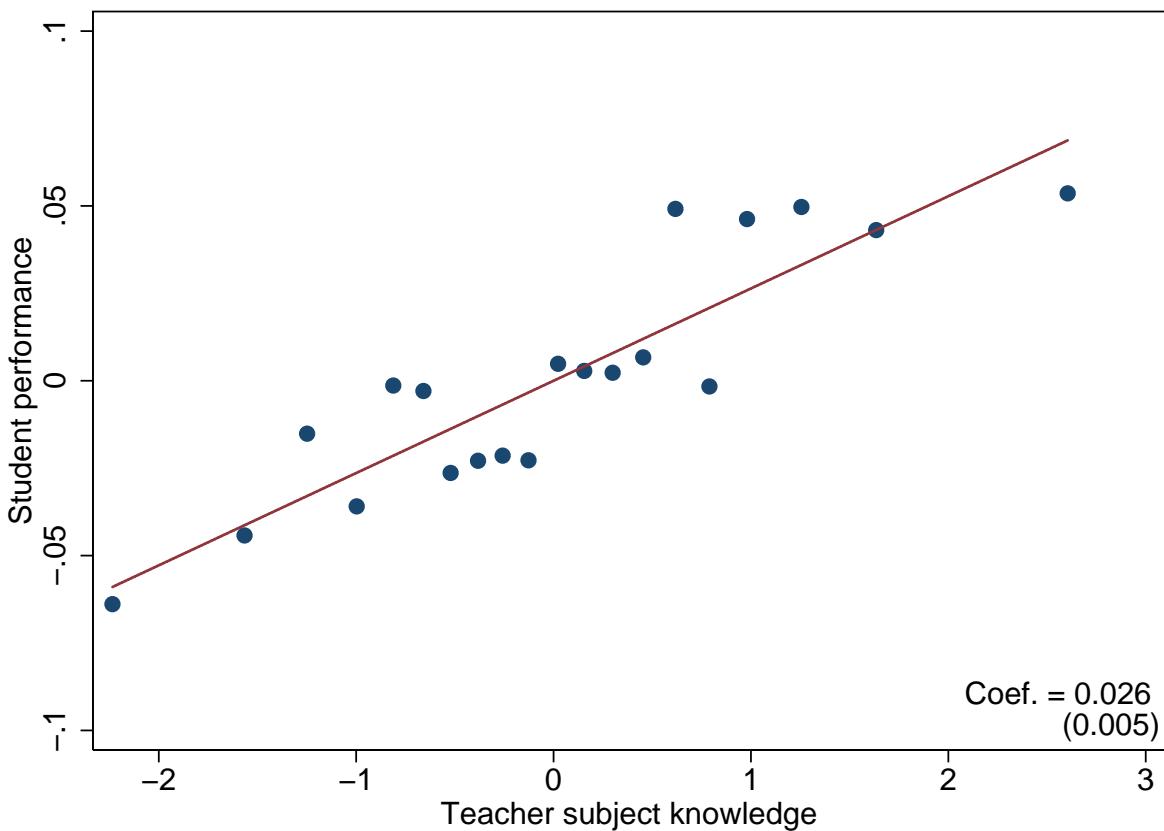
Figures and Tables

Figure 1: Potential Determinants of Cross-Country Differences in Student Performance



Notes: Solid lines fit a linear relationship between student performance and teacher subject knowledge in the top panel and between student performance and the share of college-educated teachers in the bottom panel. Share of college-educated teachers is the share of sixth-grade teachers with a college degree (based on SACMEQ data). Country abbreviations: BOT = Botswana, KEN = Kenya, LES = Lesotho, MAL = Malawi, MOZ = Mozambique, NAM = Namibia, SEY = Seychelles, SWA = Swaziland, TAN = Tanzania, UGA = Uganda, ZAM = Zambia, ZAN = Zanzibar, ZIM = Zimbabwe.

Figure 2: Effect of Teacher Subject Knowledge on Student Performance



Notes: The figure displays a binned scatterplot corresponding to the student fixed effects model in Column 3 of Table 2; see notes to Table 2 for a list of the control variables. To construct the figure, we first regressed the test score difference between math and reading of students and teachers separately on all control variables (also differences between math and reading). We then divided the teacher test score residuals into 20 ranked equal-sized groups and plotted the mean of the student test score residuals against the mean of the teacher test score residuals in each bin. The best-fit line, the coefficient, and the standard error (clustered at the school level) are calculated from regressions on the micro data.

Table 1: Ordinary Least Squares Model

| Panel A: student math performance | | | | | |
|--------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | (1) | (2) | (3) | (4) | (5) |
| Teacher math knowledge | 0.121*** (0.012) | 0.096*** (0.010) | 0.095*** (0.010) | 0.075*** (0.009) | 0.072*** (0.010) |
| Adj. R2 | 0.22 | 0.30 | 0.30 | 0.32 | 0.32 |
| Observations (students) | 74,708 | 74,708 | 74,708 | 74,708 | 74,708 |
| Clusters (schools) | 3,939 | 3,939 | 3,939 | 3,939 | 3,939 |
| Panel B: student reading performance | | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Teacher reading knowledge | 0.117*** (0.012) | 0.085*** (0.009) | 0.084*** (0.009) | 0.062*** (0.009) | 0.057*** (0.009) |
| Adj. R-squared | 0.22 | 0.35 | 0.36 | 0.39 | 0.39 |
| Observations (students) | 74,708 | 74,708 | 74,708 | 74,708 | 74,708 |
| Clusters (schools) | 3,939 | 3,939 | 3,939 | 3,939 | 3,939 |
| Control variables in Panels A + B | | | | | |
| | X | X | X | X | X |
| Country fixed effects | X | X | X | X | X |
| Socio-economic characteristics (16) | | X | X | X | X |
| Classroom characteristics (4) | | | X | X | X |
| School characteristics (5) | | | | X | X |
| Teacher characteristics (6) | | | | | X |

Notes: Least squares regressions. Dependent variable: student performance in math (Panel A) and in reading (Panel B). Student and teacher test scores are z-standardized at the individual level across countries and waves. Socioeconomic controls include three student characteristics (age, gender, repeated grade) and 13 family background measures (mother's education, father's education, number of books at home, and ten family resources). Classroom controls contain four classroom resources (availability of subject-specific textbooks, number of books in class, access to teaching guide, class size), and school resource controls include five measures of school resources and location (school facilities index (see Table 4), private school indicator, frequency of teacher absence at school, number of students in school, rural school indicator). Teacher controls include six teacher characteristics (gender, education, work experience, duration of subject-specific training, weekly teaching time, frequency of meeting parents). All regressions include imputation dummies and a dummy indicating the SACMEQ wave. Robust standard errors, adjusted for clustering at the school level, are reported in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 2: Student Fixed Effects Model

| Dependent variable: student performance | (1) | (2) | (3) |
|---|---------------------|---------------------|---------------------|
| Teacher subject knowledge | 0.025*** (0.005) | 0.025*** (0.005) | 0.026*** (0.005) |
| Student fixed effects | | X | X |
| Classroom characteristics (3) | | X | X |
| Teacher characteristics (6) | | | X |
| Observations | 149,416 | 149,416 | 149,416 |

Notes: Fixed effects estimations. Dependent variable: student performance in math and reading. Student and teacher test scores are z-standardized at the individual level across countries and waves. Compared to Table 1, among classroom characteristics, class size is excluded because it does not vary across subjects for the same student. All regressions include subject fixed effects and imputation dummies. Robust standard errors, adjusted for clustering at the school level, are reported in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 3: Robustness (Student Fixed Effects Model)

| Dependent variable: student performance | Rural schools (1) | One-classroom schools (2) | Rural & one-classroom schools (3) | School level (4) | Same-teacher sample (5) | Subject interactions (6) | Subject interactions (7) |
|--|----------------------|------------------------------|--------------------------------------|---------------------|----------------------------|-----------------------------|-----------------------------|
| Teacher subject knowledge | 0.021*** (0.006) | 0.022*** (0.008) | 0.026*** (0.009) | | 0.025*** (0.010) | 0.017*** (0.005) | 0.008* (0.004) |
| Teacher subject knowledge (school level) | | | | 0.030*** (0.005) | | | |
| Student fixed effects | X | X | X | X | X | X | X |
| Classroom characteristics (3) | X | X | X | X | X | X | X |
| Teacher characteristics (6) | X | X | X | X | n.a. | X | X |
| Subject X socio-economic characteristics (6) | | | | | | X | X |
| Subject X school characteristics (5) | | | | | | X | X |
| Subject X country fixed effects | | | | | | | X |
| Observations | 92,968 | 63,204 | 51,948 | 149,416 | 46,888 | 149,416 | 149,416 |

Notes: Fixed effects estimations. Dependent variable: student performance in math and reading. Student and teacher test scores are z-standardized at the individual level across countries and waves. In Column 1, the sample includes only schools in rural areas. In Column 2, all schools with more than one sixth-grade classroom are excluded. In Column 3, the sample includes only schools in rural areas with just one sixth-grade classroom. In Column 4, teacher test scores and all teacher characteristics are collapsed at the school level. In Column 5, the sample includes only students who are taught both math and reading by the same teacher; teacher characteristics are excluded as they do not vary within the same teacher. In Column 6, the subject indicator is interacted with students' socioeconomic characteristics (no family resources) and school characteristics (see Table 1); in Column 7, the subject indicator is additionally interacted with country fixed effects. Classroom and teacher characteristics are the same as in Table 2. All regressions include subject fixed effects and imputation dummies. Robust standard errors, adjusted for clustering at the school level, are reported in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

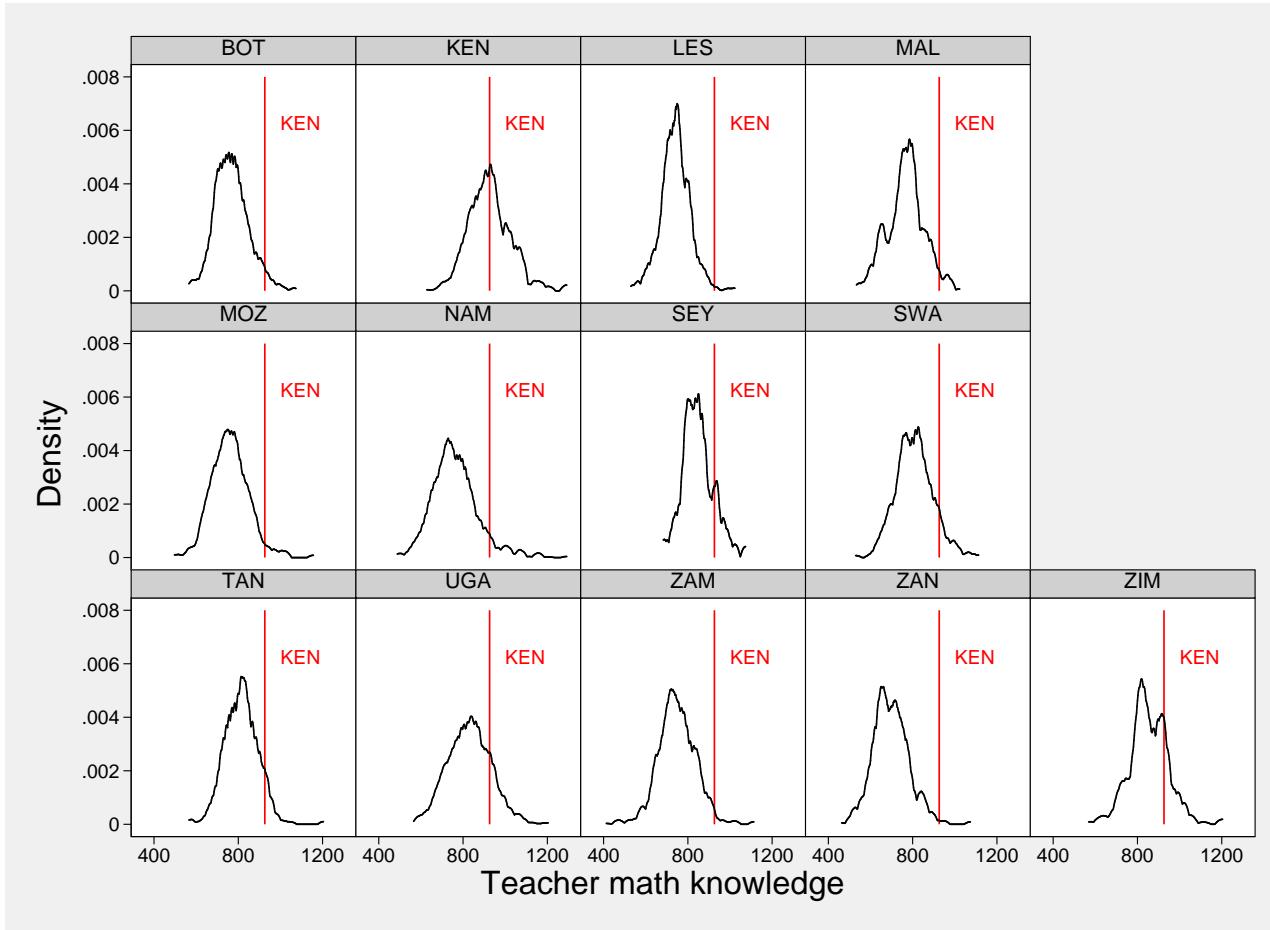
Table 4: Heterogeneity (Student Fixed Effects Model)

| Dependent variable: student performance | | | | | |
|---|---|---------------------|-------------------------|------------------------|---------------------|
| | Country's level of economic development | | Student-level resources | School-level resources | |
| | (1) | (2) | (3) | (4) | (5) |
| Teacher subject knowledge | 0.009 (0.008) | 0.011 (0.008) | 0.017** (0.007) | 0.025*** (0.005) | 0.027*** (0.005) |
| × high GDP per capita | 0.038*** (0.010) | | | | |
| × high Human Development Index | | 0.030*** (0.010) | | | |
| × textbook availability | | | 0.017** (0.007) | | |
| × school facilities (index) | | | | 0.011** (0.005) | |
| × average class size | | | | | 0.001 (0.006) |
| Textbook availability | | | 0.006 (0.010) | | |
| Student fixed effects | X | X | X | X | X |
| Classroom characteristics (3) | X | X | X | X | X |
| Teacher characteristics (6) | X | X | X | X | X |
| Observations | 143,978 | 149,416 | 149,416 | 149,416 | 146,310 |

Notes: Fixed effects estimations. Dependent variable: student performance in math and reading. *GDP per capita*: gross domestic product divided by midyear population expressed in PPP-US-\$; data from the UNESCO Institute for Statistics. The following countries have a “high” (i.e., above-median) GDP per capita: Botswana, Kenya, Namibia, Seychelles, Swaziland; data for Zimbabwe are not available. *Human Development Index*: summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable, and having a decent standard of living; data are from the African Development Bank. The following countries have a “high” Human Development Index: Botswana, Kenya, Lesotho, Namibia, Seychelles, Swaziland. We assign the same values to country-level variables in Tanzania and Zanzibar because Zanzibar is a semi-autonomous part of Tanzania. *Textbook availability*: binary variable that equals 1 if a student shares his or her subject-specific textbook with exactly one other student or has own textbook; 0 otherwise. *School facilities (index)*: counts the availability of all 31 school resources reported in SACMEQ: board, cafeteria, chairs, chalk, charts, classroom library, community hall, computer, drinking water, duplicator, electricity, fax, fence, first aid kit, garden, locker, overhead projector, photocopier, playground, radio, school library, separate office for school head, shelves, storeroom, tables, tape recorder, teacher room, telephone, TV, typewriter, and VCR. *Average class size*: average number of students per classroom in sixth grade; 3,106 student observations are missing because some principals did not report the number of sixth-grade students in their school. To facilitate interpretation of coefficient magnitudes, the resource variables in Columns 4 and 5 are z-standardized across countries and waves. The main effects of the school-level resources and country-level variables cannot be estimated because these variables do not vary across subjects. Classroom and teacher characteristics are the same as in Column 3 of Table 2. All regressions include subject fixed effects and imputation dummies. Robust standard errors, adjusted for clustering at the school level, are reported in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

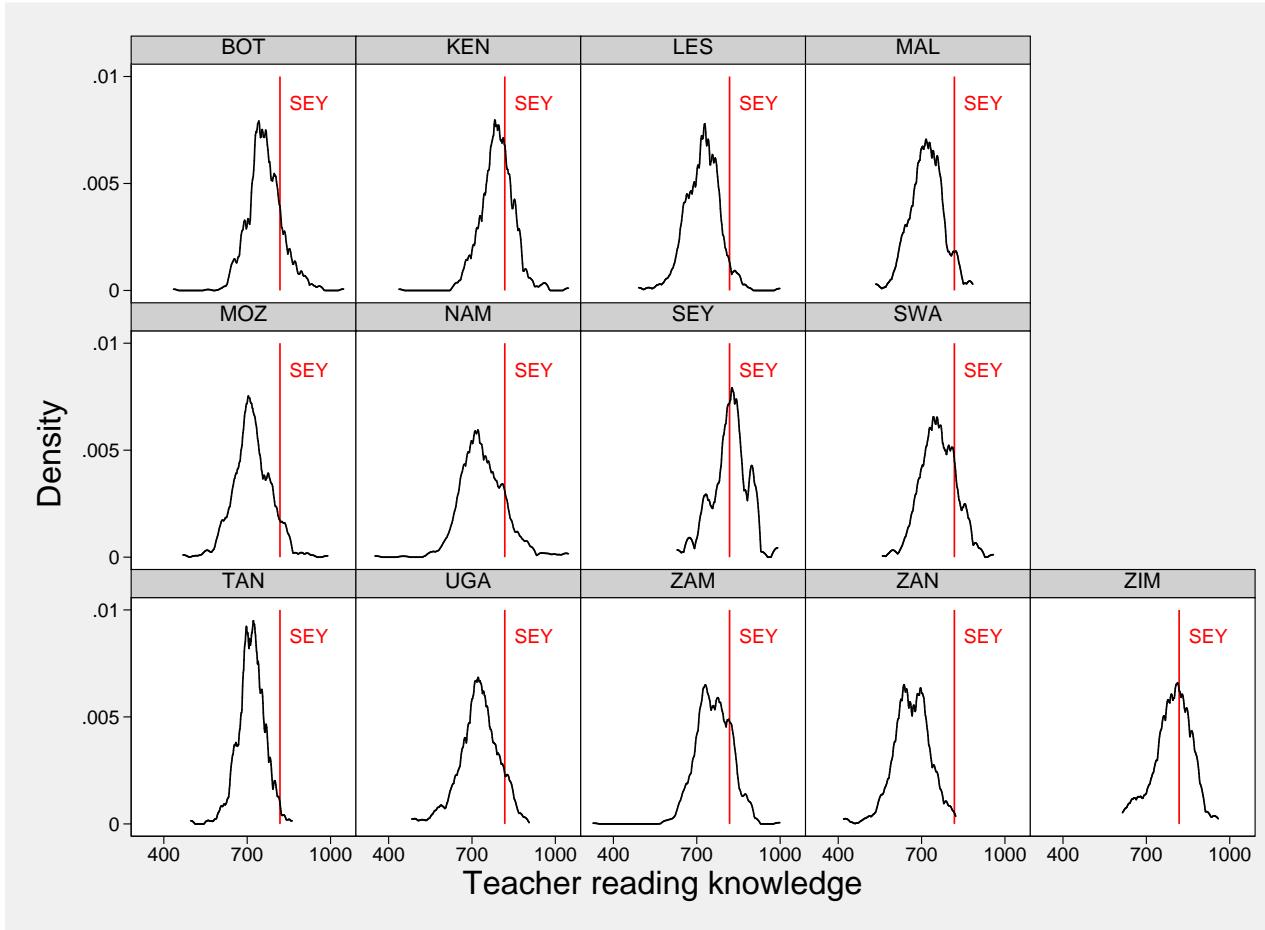
Appendix

Figure A-1: Distribution of Teacher Math Knowledge by Country



Notes: Kernel density plots of teacher math knowledge separately for each country. Vertical red lines indicate the average math knowledge of teachers in Kenya, the country in our sample with the highest average teacher math knowledge. Country abbreviations: BOT = Botswana, KEN = Kenya, LES = Lesotho, MAL = Malawi, MOZ = Mozambique, NAM = Namibia, SEY = Seychelles, SWA = Swaziland, TAN = Tanzania, UGA = Uganda, ZAM = Zambia, ZAN = Zanzibar, ZIM = Zimbabwe.

Figure A-2: Distribution of Teacher Reading Knowledge by Country



Notes: Kernel density plots of teacher reading knowledge separately for each country. Vertical red lines indicate the average reading knowledge of teachers in the Seychelles, the country in our sample with the highest average teacher reading knowledge. Country abbreviations: BOT = Botswana, KEN = Kenya, LES = Lesotho, MAL = Malawi, MOZ = Mozambique, NAM = Namibia, SEY = Seychelles, SWA = Swaziland, TAN = Tanzania, UGA = Uganda, ZAM = Zambia, ZAN = Zanzibar, ZIM = Zimbabwe.

Table A-1: Summary Statistics of Student Performance and Teacher Subject Knowledge

| | Pooled | Botswana | Kenya | Lesotho | Malawi | Mozambique | Namibia |
|---------------------|--------------|-------------|--------------|--------------|-------------|-------------|--------------|
| Students | | | | | | | |
| Math performance | 497 (88) | 517 (81) | 561 (90) | 463 (64) | 441 (61) | 504 (68) | 459 (84) |
| Reading performance | 502 (95) | 529 (94) | 548 (95) | 460 (64) | 430 (51) | 491 (72) | 482 (94) |
| # Students | 74,708 | 6,375 | 6,778 | 6,895 | 4,733 | 5,308 | 10,365 |
| Teachers | | | | | | | |
| Math knowledge | 793 (108) | 769 (82) | 926 (103) | 741 (70) | 767 (88) | 760 (90) | 763 (111) |
| Reading knowledge | 742 (74) | 764 (62) | 791 (59) | 721 (61) | 718 (60) | 718 (67) | 739 (81) |
| # Math teachers | 5,421 | 730 | 474 | 422 | 278 | 586 | 587 |
| # Reading teachers | 5,466 | 725 | 480 | 421 | 288 | 603 | 561 |
| | Seychelles | Swaziland | Tanzania | Uganda | Zambia | Zanzibar | Zimbabwe |
| Students | | | | | | | |
| Math performance | 549 (100) | 531 (64) | 542 (86) | 486 (86) | 432 (70) | 485 (66) | 513 (96) |
| Reading performance | 575 (122) | 542 (66) | 566 (89) | 477 (80) | 433 (77) | 509 (88) | 502 (100) |
| # Students | 2,820 | 6,700 | 6,455 | 6,498 | 4,745 | 4,317 | 2,719 |
| Teachers | | | | | | | |
| Math knowledge | 848 (75) | 809 (89) | 814 (81) | 830 (103) | 748 (89) | 697 (85) | 855 (93) |
| Reading knowledge | 818 (65) | 761 (63) | 715 (49) | 725 (71) | 761 (63) | 664 (66) | 794 (67) |
| # Math teachers | 91 | 336 | 397 | 355 | 534 | 362 | 269 |
| # Reading teachers | 105 | 336 | 398 | 359 | 534 | 387 | 269 |

Notes: Means and standard deviations (in parentheses) reported. The pooled sample includes 8,742 teachers in total, some of them teaching both math and reading. Statistics are based on individual-level observations.

A Measurement Error

As with any performance assessment, teacher subject knowledge in SACMEQ is likely measured with error. Measurement error in the explanatory variable may lead to a downward bias in the estimated coefficient, and this bias may be aggravated in the student fixed effects models (Angrist and Krueger, 1999, Chapter 4). In this appendix, we assess the importance of measurement error for our estimates and propose a way of correcting the corresponding attenuation bias.

We begin our analysis by assuming that teacher subject knowledge is measured with random noise. Let T_{ik}^* denote the true knowledge of student i 's teacher in subject k and let the observed teacher test score be denoted by $T_{ik} = T_{ik}^* + e_{ik}$.³⁸ Assuming classical measurement error, $E(e_{ik}) = 0$ and $Cov(T_{ik}^*, e_{ik}) = 0$. In a bivariate model, the true effect of teacher subject knowledge on student performance, y_{ik} , will then be asymptotically biased towards zero:

$$y_{ik} = \beta \lambda_k T_{ik} + \varepsilon_{ik}, \quad \text{where } \lambda_k = \frac{Var(T_{ik}^*)}{Var(T_{ik}^*) + Var(e_{ik})}. \quad (3)$$

The factor λ_k indicates how much the true effect β is attenuated and is often referred to as the reliability ratio or signal-to-noise ratio.

In a first-differenced, that is, a student fixed effects model, the attenuation bias due to measurement error is likely aggravated. Intuitively, teachers' levels of math and reading knowledge are more strongly correlated than the measurement errors in these variables, such that differencing the observed test scores decreases the signal-to-noise ratio. More formally, consider the case where the measurement errors are uncorrelated across subjects, that is, $Cov(e_{im}, e_{ir}) = Cov(T_{im}^*, e_{ir}) = Cov(e_{im}, T_{ir}^*) = 0$. In this case, the reliability ratio for the first-differenced model can be derived as (see Metzler and Woessmann, 2010):

$$\begin{aligned} \lambda_\Delta &= \frac{Var(\Delta T_i^*)}{Var(\Delta T_i^*) + Var(\Delta e_i)} \\ &= \frac{\lambda_m Var(T_{im}) + \lambda_r Var(T_{ir}) - 2Cov(T_{im}, T_{ir})}{Var(T_{im}) + Var(T_{ir}) - 2Cov(T_{im}, T_{ir})}. \end{aligned} \quad (4)$$

³⁸For conciseness, we omit classroom and school subscripts in this discussion.

Note that the only unknown quantities in Equation (A.2) are λ_m and λ_r , while the variances and covariances of teacher subject knowledge can easily be computed from the data. Therefore, if the reliability ratios of the teachers' math and reading assessments were known, we could correct our baseline estimate for measurement error by multiplying the estimated coefficient with $1/\lambda_\Delta$.

Referring to psychometric test theory, Metzler and Woessmann (2012) argue that Cronbach's α is a natural estimate for λ_k in the context of teacher subject knowledge. We compute Cronbach's α for the math and reading tests (which are not reported by SACMEQ) by using teachers' answers to all individual test items.³⁹ The estimated reliability ratios are $\widehat{\lambda}_m = 0.83$ for math and $\widehat{\lambda}_r = 0.75$ for reading. Together with $\text{Var}(T_{im}) = \text{Var}(T_{ir}) = 1$ (due to our normalization of test scores) and the estimated covariance $\widehat{\text{Cov}}(T_{im}, T_{ir}) = 0.34$, we obtain $\widehat{\lambda}_\Delta = 0.68$ as an estimate for the reliability ratio for the differenced teacher test scores. Therefore, under the assumptions set out in the previous paragraphs, multiplying our baseline coefficient by the factor $1/0.68 = 1.46$ will provide the measurement-error-corrected estimate of the effect of teacher subject knowledge on student performance. For our baseline coefficient of 0.026 SD, this implies a corrected effect of 0.039 SD.

³⁹Cronbach's α is a function of the number of test items and the covariances between all possible item pairs; see Johannes Metzler and Ludger Woessmann (2010, 2012) as well as references therein. We use Stata's `alpha` command to compute Cronbach's α for the teacher math and reading tests in SACMEQ.