POLICY RESEARCH WORKING PAPER 7956

What Do Teachers Know and Do? Does It Matter?

Evidence from Primary Schools in Africa

Tessa Bold
Deon Filmer
Gayle Martin
Ezequiel Molina
Christophe Rockmore
Brian Stacy
Jakob Svensson
Waly Wane

Abstract

School enrollment has universally increased over the past 25 years in low-income countries. However, enrolling in school does not guarantee that children learn. A large share of children in low-income countries learn little, and they complete their primary education lacking even basic reading, writing, and arithmetic skills—the so-called "learning crisis." This paper uses data from nationally representative surveys from seven Sub-Saharan African countries, representing close to 40 percent of the region's total population, to investigate possible answers to this policy failure by quantifying teacher effort, knowledge, and skills. Averaging across countries, the paper finds that students receive two hours and fifty minutes of teaching per day—or just over half the scheduled time. In addition, large shares of teachers do not master the curricula of the students they are teaching; basic pedagogical knowledge is low; and the use of good teaching practices is rare. Exploiting within-student, within-teacher variation, the analysis finds significant and large positive effects of teacher content and pedagogical knowledge on student achievement. These findings point to an urgent need for improvements in education service delivery in Sub-Saharan Africa. They also provide a lens through which the growing experimental and quasi-experimental literature on education in low-income countries can be interpreted and understood, and point to important gaps in knowledge, with implications for future research and policy design.

What Do Teachers Know and Do? Does It Matter?

Evidence from Primary Schools in Africa[*]

Tessa Bold[a], Deon Filmer[b], Gayle Martin[c], Ezequiel Molina[d], Christophe Rockmore[e], Brian Stacy[f], Jakob Svensson[g], and Waly Wane[h]

---

1. Introduction

An educated workforce is necessary for a high national standard of living. Increasing the human capital of the poor is likely one of the most effective ways to reduce poverty and increase upward economic mobility.

Over the last 25 years, school enrollment, at all levels, has increased universally, and most children in low- and middle-income countries now complete primary school. Enrolling in school, however, does not guarantee that children acquire the competencies set out in the official curriculum. In fact, a large share of children in low-income countries complete their primary education lacking even basic reading, writing, and arithmetic skills. For example, when Grade 6 students in 15 Southern and Eastern African countries were tested in 2007 as part of the Southern and Eastern Africa Consortium for Monitoring Education Quality (SACMEQ) regional assessment exercise, less than 50 percent scored beyond the level of "reading for meaning," and less than 40 percent scored beyond "basic numeracy" (Hungi et al. 2010). Among sixth-grade students who were tested as a part of the CONFEMEN Programme for the Analysis of Education Systems (PASEC) regional assessment in 10 Francophone countries in Western and Central Africa in 2015, less than 45 percent surpassed the competency level in reading or mathematics deemed "sufficient" to successfully continue their schooling (Malpel et al 2016). UNESCO has dubbed this failure the "global learning crisis" (UNESCO 2013).

A growing body of evidence suggests that teacher quality, broadly defined, is a key determinant of student learning.[1] Little is known, however, about what specific dimensions of teacher quality matter and even less about how teachers perform along these dimensions – facts we argue are crucial in order to guide both research and policy design.

This paper reports on an ongoing research program intended to help fill this void. Using data derived from direct observations, unannounced visits, and tests, from primary schools in seven Sub-Saharan African countries—Kenya, Nigeria, Mozambique, Senegal, Tanzania, Togo, and Uganda—which together represent close to 40 percent of the region's total population, we answer four questions: (1) How much time do teachers actually spend teaching? (2) Do teachers have the relevant subject content knowledge to teach basic and higher-order language and mathematics skills? (3) Do teachers have the pedagogical knowledge and skills

---

[1] See for example Rockoff (2004); Rivkin et al. (2005); Aaronson et al. (2007); Metzler and Woessmann (2012); Chetty et al. (2014); and Das and Bau (2016) for evidence based on quasi-experimental data. These findings are supported by a growing experimental literature reviewed in, for example, Kremer, Brannen, and Glennerster (2013); Glewwe and Muralidharan (2015); Bruns, Filmer and Patrinos (2014) and Evans and Popova (2016), showing that traditional educational inputs have little impact on test scores but incentivizing teacher effort and supporting specific aspects of pedagogy do.

to transfer what they know to students? (4) To what extent does teacher content and pedagogical knowledge matter?

Averaging across countries, we find that students receive about two hours and fifty minutes of teaching per day—or just over half the scheduled time. This is largely because teachers, even when in school, are not teaching. Furthermore, teachers' subject knowledge is strikingly low. Only about one in ten fourth-grade teachers master their students' language curriculum, and about a quarter of the teachers fail simple tasks (such as subtracting two-digit numbers for math teachers, or choosing the correct pronoun or conjunction to complete a sentence for language teachers). With regard to pedagogy, few teachers are able to assess children's abilities and evaluate their students' progress, and few exhibit practices that are typically associated with good teaching (e.g. regularly checking for students' understanding and giving feedback).

Exploiting the linked student-teacher data across countries, and within-teacher within-student variation, we find significant, and importantly large, positive effects of teacher content and pedagogical knowledge on student achievement.

Our findings have two main sets of implications: the first are for education systems and education policy reform; the second are for the experimental and quasi-experimental research agenda on ways to improve education quality. On the former, rapid expansions in school enrollment in developing countries have put substantial pressure on education systems' ability to provide quality education—and rapid hiring of teachers to limit student–teacher ratios may have contributed to our findings of underprepared teachers. At the same time, the high rates of teacher absenteeism from school, as well as absenteeism from the classroom while at school, point to systemic governance, accountability, and management issues. The magnitude of the problem highlighted here suggests that there is an urgent need to tackle these underlying causes of poor service delivery. Failure to do so not only represents a waste of the considerable resources that countries spend on education, but more importantly, condemns generations of students to subpar education and consequently diminished opportunities in life. The findings also suggest care in further system expansions (such as at the preschool or secondary school levels) to ensure that the same patterns are not repeated.

Our findings also provide a quantitative lens through which the growing experimental and quasi-experimental literature on education can be interpreted and understood. Over the last 15 years, more than 200 randomized controlled trials have been conducted in the area of education (Evans and Popova 2016). However, the literature has yet to converge to a consensus among researchers about the most effective ways to increase the quality of primary education, as recent

systematic reviews demonstrate.[1] In particular, our findings help in understanding both the effect sizes of interventions shown to raise students' test scores, and the reasons why some well-intentioned policy experiments have not significantly impacted learning outcomes. At the core is the interdependence between teacher effort, ability, and skills in generating high quality education. For example, interventions that focus on increasing one component—say, low teacher effort—may not have as high an impact as expected if teachers' knowledge of the subject they teach is too low. For the same reason, formal teacher education and training do not explain much of the variation in test scores across students in our data, even though their subject and pedagogical knowledge are superior, because those with more formal training tend to teach less.

Most importantly, we argue that the stylized facts presented here can help guide the next generation of policy experiments and reforms. For example, we document substantial shortcomings in teacher knowledge and skills, and show, using quasi-experimental variation, that teacher subject knowledge and teacher pedagogy knowledge and skills are important determinants of student test scores. As of yet, however, we know relatively little about effective ways to raise these aspects of teacher competencies.

We proceed by first providing a brief background of the research program (Service Delivery Indicators) followed by a short description of measured student learning in the seven countries surveyed (sections 2 and 3). The subsequent three sections (4, 5, and 6) aim to answer the main questions outlined above: How much do teachers teach? What do teachers know? How well do teachers teach? Sections 7 and 8 exploit the data to account for the variation in teacher effort, knowledge, and skills across schools, and to account for and explain the variation in student learning. Section 9 compares teachers in public and private schools. Finally, in a concluding

---

[1] Kremer, Brannen, and Glennerster (2013), drawing on a set of RCT studies, argue that interventions that match teaching to student learning levels, contract teachers, and interventions improving access to schooling are the most effective. Krishnaratne, White, and Carpenter (2013) – a meta-analysis of 69 RCT and quasi-experimental studies – argue that the most compelling evidence of what works is computer-assisted learning tools. McEwan (2014) – a meta-analysis of 77 RCT – finds the largest effects for interventions involving computer-assisted learning. Murnane and Ganimian (2014) – a narrative review drawing on 115 RCT and quasi-experimental studies – conclude that the strongest evidence (unconditionally) of impact comes from studies providing information about school quality and returns to schooling. Conn (2014) – a meta-analysis based on 56 studies conducted in Sub-Saharan Africa – finds that pedagogical interventions (changes in instructional techniques) have the highest effect size on achievement outcomes, and Glewwe and Muralidharan (2015) – employing a voting counting approach based on 118 studies (of which 80 RCTs) – conclude that teaching at the right level (remedial programs), and teacher performance and accountability interventions are the most promising. Evans and Popova (2015) provide a review of these reviews to assess which findings are consistent, and which divergent, across these studies and conclude that there are three areas where interventions tend to be most consistently impactful: pedagogical interventions (including computer-assisted learning) that tailor teaching to student skills; repeated teacher training interventions, often linked to another pedagogical intervention; and improving accountability through contracts or performance incentives, at least in certain contexts.

section, we briefly discuss implications for policy and research. There, we highlight the scale of the problems facing policy makers and other stakeholders, discuss how the findings can help improve our understanding of the growing experimental literature on education in developing countries, and point to avenues for future research.

2. Measuring teacher effort, knowledge, and skills: The Service Delivery Indicators

The Service Delivery Indicators (SDI)—an ongoing Africa-wide program with the aim of collecting informative and standardized measures of what primary teachers know, what they do, and what they have to work with—grew out of concern about poor learning outcomes observed in various student tests as well as evident shortcomings in fast-expanding systems of education. This policy failure is evident throughout the education service delivery chain, but it is most clearly (and perhaps most damagingly) manifested at the school level.

The delivery of education in many low-income countries is characterized by centralized, but typically weak, state control and often low-capacity, locally governed institutions for education provision. At the same time, the institutional incentives for performance are largely missing, with both career progression and financial rewards delinked from performance. Hiring, salaries, and promotions are largely determined by teachers' seniority and initial educational qualifications, and are unrelated to effort or performance. In most settings, parents have little influence on how teachers are hired or schools are managed, and the various state and local authorities provide limited technical support or supervision.

Teacher salaries account for the largest single item in education expenditure, in countries at all income levels. In Sub-Saharan Africa, salaries for teachers and education officials account for more than 70 percent of the expenditure in education (UIS/ISU 2013) and approximately 12 percent of total government expenditure. Teachers in Sub-Saharan African countries earn on the order of 4 times average GDP per capita (UIS 2011) whereas in high-income countries the ratio is closer to 1.5 (Bruns and Luque 2014).

In Africa, the public sector is the dominant actor in primary education. However, while public spending on education has increased in the last decade, so has the number of private schools. Recent data suggest that private schools—both informal and formal—account for around 20 percent of total primary school enrollment in low-income countries (Baum et al. 2014).

The SDI program has to date been implemented in Tanzania and Senegal in 2010 (Bold et al, 2011), Kenya (2012), Mozambique (2014), Nigeria (2013), Tanzania (2010, 2014), Togo (2013), and Uganda (2013), bringing the total to eight surveys in seven countries. The surveys

were national in scope, with the exception of Nigeria, where surveys representative at the state level were implemented in four states (Anambra, Bauchi, Ekiti, and Niger).

Representative surveys of between 150 and 760 schools were implemented in each country or Nigerian state using a multistage, cluster-sampling design. Primary schools with at least one fourth-grade class formed the sampling frame. The samples were designed to provide representative estimates for teacher effort, knowledge, and skills in public primary schools, broken down by urban and rural location. For five of the six non-pilot surveys, representative data were also collected for private primary schools. Across the eight surveys, the SDI collected data on 2,600 schools, over 21,000 teachers and 24,000 students in Sub-Saharan Africa.[2]

The surveys collected a broad set of school, teacher, and student specific information, with an approach that relies as much as possible on direct observation rather than on respondent reports. Data were collected through visual inspections of fourth-grade classrooms and the school premises, direct physical verification of teacher presence by unannounced visits, and teacher and student tests. We focus here on the data on teacher behavior and knowledge.[3]

3. The starting point: Learning outcomes in primary schools in Africa

In the last decade, the major International Association for the Evaluation of Education Achievement (IEA) and OECD testing programs have expanded dramatically, with more than 100 participating countries in at least one of these assessments in 2012 (Hanushek and Woessmann 2015). However, only one Sub-Saharan African country (Botswana) participated in the last IEA mathematics tests at the primary level, and only three countries participated at the secondary level (Botswana, Ghana, and South Africa).[4] Average test results at the secondary level for these three countries suggest *average* test scores below the lowest *5th percentile* score in the US. A significant share of students, however, performed worse than chance (based on multiple choice items), suggesting that the tests are potentially unreliable at pinpointing performance at such low levels (TIMSS 2011).

As part of the SDI assessment of teacher behavior and knowledge, fourth-grade students in sampled schools were assessed in basic reading, writing, and arithmetic skills. While other

---

[2] More details on the sample are available in an Appendix available from the authors upon request.
[3] More information on SDI can be obtained at www.worldbank.org/sdi.
[4] The Trends in International Mathematics and Science Study (TIMMS) by the International Association for the Evaluation of Education Achievement (IEA) test Grade 4 (primary) and Grade 8 (secondary) students on common elements of primary and secondary school curricula. In Botswana, Grade 6 students were assessed, as it was deemed too difficult for fourth-grade students to take the TIMSS fourth-grade assessment. Similarly, in Botswana and South Africa, ninth-grade students were assessed with the TIMSS eighth-grade assessment. No Sub-Saharan African country participated in the testing program run by OECD (Programme for International Student Assessment, PISA).

testing programs exist in Sub-Saharan Africa, including SACMEQ, PASEC, and Uwezo (which uses volunteers to assess the basic literacy and numeracy of children in a national sample), the main advantage of the SDI assessment is that it is possible to link student achievement to teacher characteristics and observations.

The SDI student test was designed as a one-on-one evaluation, with enumerators reading instructions aloud to students in their mother tongue. This was done in order to build up a differentiated picture of students' cognitive skills; i.e. oral one-to-one testing allows one to test whether a child can solve a mathematics problem even when his/her reading ability is so low that he/she would not be able to attempt the problem independently. The language test, which evaluated ability in the language of instruction English, French, or Portuguese, ranged from simple tasks that tested letter and word recognition to a more challenging reading comprehension test.[5] The mathematics test ranged in difficulty from recognizing and ordering numbers, to the addition of one- to three-digit numbers, to the subtraction of one- and two-digit numbers, and to the multiplication and division of single-digit numbers. We focus on young students here because cognitive skills are more malleable at this age and because of the mounting evidence on the importance of education investments at young ages.

As evident, after three completed years of primary public schooling, many students still lack even basic literacy and numeracy skills (Table 1 summarizes the findings by listing results for a handful of items covered in the language and mathematics assessment). For example, almost half of the students assessed could not read a simple word, ranging from 79 percent who could not in Portuguese in Mozambique to 18 percent in English in Kenya. A majority of the students (71 percent overall, ranging from 29 to 89 percent) could not read all the words in a basic sentence. Furthermore, the vast majority of students could not read a simple paragraph and infer meaning from it by answering three questions of comprehension (89 percent overall, ranging from 74 to 98 percent).

In mathematics, 12 percent of students could not recognize numbers, and 55 percent could not order numbers ranging from single to triple digits; 24 percent could not add single digits; and 34 percent could not subtract single digits. 44 percent could not add double digits, and 70 percent could not subtract double digits.

While some of the variation in literacy scores across countries is undoubtedly related to variation in students' immersion in the language of assessment across countries, the same caveat

---

[5] In Tanzania (2014), a randomly selected half of the students were administered the English version of the test, and the other half were administered the Swahili version. For consistency in this analysis, we use only the students that were tested in English.

does not apply to their numeracy skills. Importantly, it is difficult to argue that such large differences at the country level should be related to differences in students' innate ability rather than to systematic differences in the quality of service delivery, which we explore in the next two sections. This is also borne out by the fact that the coefficient of variation of students' non-verbal reasoning skills, a measure designed to capture differences in innate ability, is one-third of the coefficient of variation of students' math scores when comparing data at the country level.[6]

Compared to rural India, where comparable data exist, the results suggest that learning outcomes across Sub-Saharan Africa are lower in language and similar for mathematics. A recent evaluation from India, for example, shows that 38 percent of children in third grade in public schools could not read simple words, and less than 27 percent could master double digit subtraction (ASER 2013).[7]

In the next sections we diagnose three teacher-related factors that underpin these poor learning outcomes: the time spent teaching, teacher knowledge, and teacher practices and behaviors.

4. How much do teachers teach?

Being present in the classroom is a condition sine qua non for teachers to exert effort at teaching. To measure the time teachers spend teaching, SDI extended the approach in Chaudhury et al. (2006). In each school, during a first announced visit, up to 10 teachers were randomly selected from the teacher roster. At least two teaching days after the initial survey, an unannounced visit was conducted, during which the enumerators were asked to identify whether the selected teachers were in the school, and if so, if they were in class teaching during a time when they were scheduled to be. Both assessments were based on directly observing the teachers and their whereabouts.

Averaging across countries, 44 percent of teachers were absent from class, either because they were absent from school, or because they were in the school, but not in the classroom.[8] In three of the eight surveys, more than half of the teachers were absent from the classroom, and

---

[6] Details of the non-verbal measure are available in Ozier (2017). More details on this analysis are available in an Appendix available from the authors upon request.

[7] For comparison, 92 percent of Grade 3 students in Sweden are deemed to master these basic as well as much more sophisticated number operations (Stockholms Universitet, 2015).

[8] Note that we distinguish between (1) absence from school, (2) absence from class conditional on being in school, and (3) absence from class. We often refer to the last of these because it combines the effects of the first two, and captures teacher absenteeism as experienced by the students.

only in one country—Nigeria—do we observe average absence below 30 percent. Being absent from school is about as common as being present in the school but absent from class. The rank correlation coefficient between the two measures is less than 0.5 at the country level, making the school absence rate at best a partial measure of teacher effort (Table 2). This is most starkly illustrated in the cases of Kenya and Tanzania (2014), both of which have relatively low school absence rates (15 percent) but relatively high classroom absence rates conditional on being in school (38–39 percent).

When a large share of teachers is not teaching, unsurprisingly, a large share of classrooms will be occupied by only students. Consistent with the absenteeism findings discussed above, we find, averaging across countries, that one-third of the classrooms were "orphaned" classrooms, where students are present but there is no teacher.

Over time in these countries, the absenteeism rates appear remarkably stable. Chaudhury et al. (2006) estimated a school absence rate of 27 percent in Uganda in 2002–03, which compares to our measure of 30 percent in 2013. Similarly, while absence from school fell by a third in Tanzania between 2010 and 2014, this was largely offset by an increase in absence from the classroom conditional on being in school; the net result being a small decline in absence from class between the two surveys.

What do these results imply for the amount of instruction time that students receive? To answer this, the surveys first recorded the scheduled time of a teaching day—after break times—according to school records. Averaged across schools and countries, this comes to 5 hours and 27 minutes. We then multiply this number by the proportion of teachers absent from class. If 10 teachers are supposed to teach 5 hours and 27 minutes per day, yet 4 teachers are absent from either the school or the classroom at any one time, then the scheduled teaching time is reduced to 3 hours and 16 minutes.

Moreover, even when in the classroom, teachers may not necessarily be teaching. We address this by using the results from the classroom observation carried out as a part of the survey – an observation schedule based on recording a minute-by-minute snapshot of what the teacher was doing, for a randomly selected fourth-grade mathematics or language class.[9]

The percentage of the lesson lost to non-teaching activities varied from 18 percent in Nigeria, the country with the lowest classroom absence rate, to 3 percent in Uganda, the country

---

[9] This approach likely provides an upper bound on the time devoted to teaching. In the two pilot countries, Tanzania and Senegal, data on teacher behavior were collected while observing the teacher both inside and from outside the classroom. We find that time spent teaching when in class is about 30 percent lower in Senegal and about 10 percent lower in Tanzania if the enumerator stands outside the classroom rather than inside.

with the highest classroom absence rate. We then combine the absence-adjusted teaching time with the proportion of classroom time devoted to actual teaching activities to estimate instruction time as experienced by students.

Students are taught, on average, 2 hours and 49 minutes per day, or roughly half of the scheduled time (Table 2). Estimated instruction time varies from 3 hours and 16 minutes in Tanzania to 1 hour and 43 minutes in Mozambique.[10] Only about 10 percent of the schools provide more than 5 hours of teaching per day. About the same share provide no teaching (because none of the 10 randomly selected teachers was found in the classroom). More than a quarter of schools teach less than 2 hours, and half the schools teach less than 3 hours.

5. What do teachers know?

For teachers to be effective, they must have the knowledge necessary for good teaching. A first necessary requirement is that they exhibit a clear understanding of the subject they teach (subject content knowledge).

To measure the content knowledge of primary school teachers, and specifically those teaching lower primary, all language and mathematics teachers teaching Grade 4 in the current year, or Grade 3 in the previous year, were assessed in each school. On average, 5 teachers were tested in each school.

Teachers were asked to mark (or "grade") mock student tests in language and in mathematics, based on the fourth-grade curriculum.[11] In contrast to other approaches, where teachers take exams, this method of assessment aimed to assess teachers in a way that was consistent with their normal activities—namely, marking student work—and to recognize teachers as professionals. In the analysis, we only assess the language knowledge of those teachers who teach language, and the mathematics knowledge of those teachers who teach mathematics. Importantly, in interpreting the results one should bear in mind that all questions on the teacher test were based on common items taken from the primary curricula of each country, while 90 percent of teachers tested have secondary or higher education.

---

[10] For comparison, primary schools in the UK schedule on average 4 hours of teaching a day (https://eal.britishcouncil.org/learners/school-life-england), while in Germany the typical teaching time would be 3 hours and 45 minutes (http://www.bildungsxperten.net/wissen/was-ist-die-primarstufe).
[11] The subject test was designed by experts in international pedagogy and validated against 13 Sub-Saharan African primary curricula (Botswana, Ethiopia, Gambia, Kenya, Madagascar, Mauritius, Namibia, Nigeria, Rwanda, Seychelles, South Africa, Tanzania, and Uganda). See Johnson, Cunningham and Dowling (2012) for details.

We start by assessing whether teachers master their students' curriculum in language. For this, we focus on tasks that were common across the student and the teacher test—specifically, spelling and simple grammar exercises. To allow for some margin of error, we count a teacher as mastering the student curriculum if he or she marked 80 percent or more of the spelling and grammar questions correctly.

Two-thirds of teachers make it over that very low bar, though there is wide variation across countries (Table 3). While over 90 percent of teachers in Kenya and Uganda master the knowledge that their students are supposed to learn, only a quarter of Nigerian teachers do.

Possessing knowledge equivalent to the fourth grade curriculum is, of course, not sufficient to teach language in lower primary, because language teaching is "monolithic." That is to say, teaching a student how to compose even a simple text requires knowledge that goes well beyond the curriculum.

We therefore deem a language teacher in Grade 4 to have minimum subject content knowledge if he or she can confidently correct children's work in such aspects of literacy as reading comprehension, vocabulary, and formal correctness (grammar, spelling, syntax, and punctuation), all of which are competencies a teacher in lower primary would routinely be required to teach. To this end, the language test contained (in addition to the spelling and grammar exercises) Cloze passages to assess vocabulary and reading comprehension, and a letter written to a friend describing the student's school, which the teacher had to mark and correct.[12]

We formally define "minimum knowledge in language" as marking at least 80 percent of the items on the language test correctly. Only 7 percent of the language teachers meet this minimum subject knowledge, with the level uniformly low across the eight countries: in Kenya, 34 percent of language teachers have minimum subject knowledge, and no teachers in Togo, Mozambique, Tanzania (both in 2010), or Nigeria meet the threshold.[13]

Which areas of language teaching specifically need improvement? First, some teachers are weak in all areas of the curriculum: 14 percent could not spell a simple word ("traffic," for example), and a similar share could not correctly answer a grammar exercise that asked them to identify the option, out of three, that would complete a sentence such as "[_____] [*Who*, *How much*, *How many*] *oranges do you have?*" Second, most teachers struggled with those

---

[12] Cloze passages are passages that intentionally have missing words to be filled by the respondent.

[13] The conclusion that many language teachers have low subject knowledge does not change if one were to drop the requirement to scoring 60% or more on the language portion of the test, or if one defined the indicator over different combinations of items on the language test (further details are in Appendix available from the authors upon request).

tasks that required at least some knowledge beyond the lower primary curriculum to mark. Less than half correctly marked the Cloze passage which included "student" responses such as "*[Where] do I have to go to the market?*" (the correct answer being "*Why*" or *"When")*. Teachers corrected only a quarter of the spelling, grammar, syntax, and punctuation mistakes in a child's letter that included segments such as "*I went to tell you that my new school is better the oldone I have a lot of thing to tell you about my new school in Dar es Salaam*."

In mathematics, we classify a teacher as having minimum subject content knowledge if he or she can accurately correct children's work in such aspects of numeracy as manipulating numbers using whole number operations. This requirement amounts to correctly scoring 80 percent or more of the questions on the lower primary portion of the mathematics test. In essence, the test thus measures whether the math teacher masters his or her students' curriculum, allowing for a 20 percentage-point margin of error. Fewer than 70 percent of mathematics teachers have minimum knowledge according to this definition (Table 3), although there is again wide variation across countries. In Senegal, 75 percent of mathematics teachers are deemed to have minimum knowledge, but in Nigeria, only half the mathematics teachers meet this standard (Table 3).

Overall, 77 percent of teachers could subtract double digits, but only around 60 percent could do so in Nigeria (Table 4). Similarly, 68 percent of teachers could multiply double digits, but less than 60 percent of math teachers in Mozambique could do this. When it comes to understanding and solving a simple math story problem, half managed to do so, but only about 30 percent gave the correct answer in Togo. One in ten teachers could not answer the simplest question, adding two double digit numbers.

That the two measures of teacher knowledge (i.e. knowing the students' curriculum and minimum knowledge for teaching) coincide for mathematics teaching—but not for language— is a consequence of the subject's modular nature. In other words, it is possible, in principle, to teach fourth graders how to divide two numbers without having a deeper knowledge of algebra. As a consequence, the number of teachers considered to master their students' curriculum is very similar for language and mathematics, while there is a large difference in the number of teachers considered to have minimum knowledge for teaching between the two subjects.

Of course, we would expect a competent math teacher to have knowledge beyond that of his or her students, and the mathematics test, therefore, also included questions one would only encounter in upper primary school. Many mathematics teachers struggled with these tasks: only a minority of teachers, and in some countries almost none, could interpret information in a Venn diagram and/or a graph (Table 4 and Table A6). As we will see below, this low competence in

interpreting data has implications for teachers' ability to monitor their students' progress. Finally, only a few teachers (14 percent) could solve a more advanced math story problem, and only one-third could solve a logic puzzle.

Comparing student knowledge (Table 1) and teacher content knowledge (Table 4) findings provides a preview to the analysis of student learning results in Section 8. Pooling the data across the seven countries we find, for example, that almost one in three students master subtraction of double digit numbers, but almost one in three teachers do not. Importantly, the large majority of students demonstrating knowledge of double digit subtraction are taught by teachers that also master this task.

6. How well do teachers teach?

Knowing one's subject is a necessary, but not sufficient, condition for good teaching. Teachers must also know how to translate their subject knowledge into effective pedagogy and then apply this in the classroom. There is broad agreement that for teaching to be effective, lessons must be well-designed and well-structured. Teachers must also know how to assess student capabilities and react appropriately, for example, by asking questions that require various types of responses and by giving feedback on those responses, commonly referred to as knowledge of the context of learning (see Johnson, 2011, Coe, Aloisi, Higgins and Major, 2014, Ko and Sammons, 2013, Mujis et al., 2014, Vieluf et al., 2012).[14]

Related to this, a recent review (Mujis et. al 2014) identifies the following elements as key when it comes to behavior in the classroom: (i) designing and structuring lessons, and in particular, introducing topics and learning outcomes at the start of the lesson and reviewing them at the end; (ii) frequently checking for student understanding by asking questions, and allowing time for students to review and practice what they learned, either individually or in groups; (iii) varying the cognitive level of questions by mixing lower and higher order questions; and (iv) providing substantive feedback to students by acknowledging correct answers in a positive fashion and correcting wrong answers, as skills and practices that are consistently associated with gains in student learning.

We broadly follow the education literature and first measure teachers' pedagogical knowledge; then, we examine how well teachers can assess students and monitor their progress; and finally, we gauge the extent to which teachers use "good practices" in the classroom.

---

[14] The teacher test and the classroom observation were both designed to measure attributes about which there tends to be broad agreement as to the relevance for effective teaching.

To measure general pedagogical knowledge, the survey asked teachers to read and extract information from a factual text (general content knowledge) and to prepare a lesson, including learning aims and outcomes based on their reading (pedagogical content knowledge).[15] Both these tasks are consistent with professional tasks normally expected of teachers, and we therefore consider a teacher to have minimum general pedagogy knowledge if he or she scores 80 percent or more on this portion of the test.

To measure teachers' ability to assess students' learning and give feedback (which we shorten here to "assessing students"), they were asked to prepare questions that required students to recall what was learned (lower order) and questions that asked students to apply the material to new contexts (higher order) on the basis of their reading of the factual text. In a second task, teachers were asked to use a marking scheme to give feedback on strengths and weaknesses in students' writing and to distinguish weak and strong learners. In a third task, teachers were provided with a list of students' grades; they were then asked to turn the raw scores into averages and to comment on the learning progression of individuals and groups of students with the help of a bar chart. We define a teacher as having "minimum knowledge in assessing students" if he or she could answer 80 percent of the items in the three tasks correctly.

To quantify teaching practices, observers used a modified Stallings (1980) classroom observation snapshot module to record which activities, from a pre-determined list, the teacher performed during each minute of a lesson. Across the six countries for which this analysis can be done, only 10 percent of teachers reached the threshold for minimum general pedagogy knowledge (Table 3).[16] In four countries, fewer than 5 percent of teachers met the threshold. While teachers could usually read and understand the factual text, they were typically not able to translate this information into teaching, as they struggled to formulate lesson aims and learning outcomes based on their reading (average score of 22 percent on this task, Table 4).

Poor knowledge of general pedagogy translated into poor skills in the classroom: fewer than 50 percent of the teachers explained the topic of the lesson at the start and summarized what was learned at the end, and almost 40 percent of lessons seemed unplanned to the observers (Table 5).

---

[15] The parts of the test assessing pedagogical knowledge were designed to be consistent with what Sub-Saharan African countries might reasonably expect of their teachers and based on a review of policy documents from developing and developed countries on teacher standards that set out minimum requirements for teachers' professional practice and conduct (Johnson, Cunningham and Dowling 2011).

[16] The pedagogy test in Senegal and Tanzania (2010) was structured differently, so we omit these two countries from the indicators here. In Tanzania (2014), the pedagogy test was conducted in Swahili.

As with general pedagogical knowledge, few teachers demonstrated an ability to assess student learning and respond to that assessment. Very few could formulate questions that checked basic understanding based on what they had read, and fewer still could formulate a question that asked students to apply what they had learned to other contexts (Table 4).

During their lessons, most teachers (80 percent) asked questions that required students to recall information or to practice what was learned, but only just over 50 percent asked questions that required higher order skills and encouraged students to apply what was learned to different contexts and be creative (Table 5). Overall, only 31 percent of teachers mixed lower and higher order questions in their class—ranging from 14 percent of teachers in Mozambique to 44 percent of teachers in Uganda.

Seventeen percent of teachers could give feedback on strengths and weaknesses in students' writing using a marking scheme (Table 4)—ranging from 10 percent in Mozambique to 30 percent in Kenya. Furthermore, just over 12 percent could monitor and comment on the learning progression of students—ranging from 5 percent in Togo to 36 percent in Nigeria.

This was mirrored in the classroom: in response to students' answers, less than half the teachers gave positive feedback and corrected mistakes without scolding students, with a low of 17 percent in Mozambique and a high of 66 percent in Uganda.

In summary, general pedagogical knowledge and the ability to assess students' learning and respond to that assessment is poor across the seven countries. While many teachers deploy some of the teaching practices identified in the literature as promoting learning, very few (roughly one in ten) apply the full set of beneficial skills in their lessons.

7. Who are the 'good' teachers?

Teacher quality, as measured by their knowledge, attendance and practices in the classroom, is in general low in the seven countries we have surveyed. But we also note significant variation. We next turn to exploring which observed teacher characteristics are correlated with better quality teaching in government primary schools. In other words, which teachers are absent less often, are more knowledgeable and have better skills in the classroom.

To this end, we pool the data at the teacher level across all countries and link teacher classroom absence, subject knowledge, pedagogy knowledge, and an index of classroom practices to the following teacher characteristics: gender, age, experience, education and position in the school, where we distinguish between head teachers, contract teachers and regular teachers. Specifically, absence from classroom is a dummy set to 1 if the teacher is not found in the classroom and zero otherwise, the knowledge variables measure the percentage of

correct questions on the teacher test (ranging from zero to 1). For the index of classroom practices, we use the average of the four skills presented in Table 5, i.e. structuring, planning, asking lower and higher order questions and giving feedback, which again ranges from zero to 1.

In Table 6, we first present a set of partial correlations, where we regress each dimension of quality on the characteristic in question.[17] Controlling only for country and month and day of survey fixed effects, we see that women are significantly less likely to be absent from the classroom than men (8 percentage points, or an almost 20 percent lower absence rate than the average), while a teacher's experience and age are not correlated with absence. Teachers who are more highly educated (have either a university or an education degree), on the other hand, are more likely to be absent (3 points, or 7 percent higher absence rate than the average). In contrast, and consistent with findings from several randomized controlled trials, contract teachers, who tend to earn less and have less secure tenure than regular teachers, are less likely to be absent than regular teachers (7 points, or 15 percent lower absence rate compared to the average).[18] As discussed in Chaudhury et al. (2006), a teacher's salary and other benefits (such as job security) are related mainly to the teacher's experience and educational background. By implication, there is a negative relationship between a teacher's salary and her effort. Theoretically, teachers from the local area might be expected to be absent less often, because they care more about their students or are more accountable to the community; or absent more often, perhaps, because they have more outside opportunities in the local economy. We find the latter to be the case: teachers born in the same district as the school are significantly more likely to be absent (around 3 points). Strikingly, we find that head teachers have a 20-percentage-point higher classroom absence rate than regular teachers.[19] Importantly, in schools where the head teacher is absent, regular teachers are almost 20 percentage points more likely to be absent as well (results not reported).

Many of the teacher characteristics are of course correlated. For example, head teachers tend to have higher education and more than eight out of ten are male, while contract teachers are more likely to be women and come from the local area. To account for these correlations, we

---

[17] Additional descriptive statistics on teachers' characteristics are in an Appendix available from the authors upon request.

[18] See, for example, Duflo et al., 2015; Muralidharan and Sundararaman, 2013, Bold et al., 2016.

[19] That head teachers have higher classroom absence rates is perhaps not surprising. We note however, that their school absence rates are also significantly higher and that head teachers were only included in the SDI sample if they are also actively teaching in the school. (For absence rates by position in the school see the Appendix available from the authors upon request.)

also present a multivariate regression of the teacher's classroom absence rate on all these characteristics, again controlling for country and day and month of survey fixed effects. The coefficients change little, with the exception of the effect of contract teacher status on classroom absence, which is halved relative to the bivariate regression and the significant coefficient on experience, which is small, however (for each five years of experience, the classroom absence rate is 1 point higher).

When it comes to teacher subject and pedagogical knowledge, teachers with a university degree and/or education degree perform significantly better on the test (4 points higher scores on both parts) and also display better skills in the classroom (7 points higher score). These conclusions remain true even when controlling for other teacher characteristics. Still, while significant, a degree in education or a university degree is only associated with a one-fifth standard deviation increase in content knowledge. Younger teachers score better on the pedagogy portion of the test, perhaps because their teacher training is more recent. The effect of age remains significant also when controlling for other teacher characteristics, but is small. Finally, we note that neither women nor contract teachers, both of whom had significantly higher presence in the classroom, have worse content or pedagogy knowledge, and when it comes to classroom skills, women perform significantly better (7 points, or 20 percent increase compared to the average).

While one should be careful in not over-interpreting correlations and summary statistics, the findings reported in Table 6 suggest that, in terms of orders of magnitude, roughly the same academic student achievement is achieved by a female contract teacher as by a male teacher with an advanced degree—and the former costs significantly less in terms of public funds. Importantly, however, switching the latter for the former would not significantly address what we argue is the core problem: students learn too little from primary schooling.

8. To what extent does teacher quality matter?

So far we have focused on using the data to establish descriptive facts. The linked student-teacher data, however, also allow us to go one step further and to examine the extent to which teacher quality matters for student achievement. In other words, we ask whether, and to what extent, students' academic performance would increase if they were taught by better teachers, in particular those with sufficient content and pedagogical knowledge.

To assess the effect of teacher quality on student achievement we face several challenges. First, the four dimensions of teacher quality (all linked to individual student measures of test

scores) are available at different levels of aggregation. Measured classroom skills vary only across schools since only one teacher per school was observed. Measured teacher absence and pedagogy knowledge vary across teachers, though one could argue that some items in the pedagogy section are subject-specific (and therefore vary across teacher-subjects). Finally, teacher content knowledge varies across teacher-subjects.

Second, the measures of teacher quality are based on tests and one unannounced visit. They therefore provide imperfect measures of true teacher quality, which implies that the key explanatory variables are measured with error. In fact, for the teacher absence rate, we believe it is more appropriate to aggregate the result to the school level, as measuring individual teacher absence based on one observation only is too noisy.[20] Even after this adjustment, substantial measurement error probably remains.

Third, to estimate causal effects, we need to account properly for non-random sorting of students and teachers and omitted student and teacher characteristics. Importantly, we are constrained in our ability to do so by the different levels of aggregation for the quality measures. As a result, credible identification of causal effects is possible for some dimensions of quality (teacher subject knowledge), but is more challenging for others (teaching skills).

We now discuss how we deal with each of these challenges. Let y_{ijs} denote the academic achievement of student i in school j in subject $s = \{1,2\}$. Our general specification relates academic achievement to a set of observable and unobservable student, school, and teacher specific variables according to

$$y_{ijs} = \beta_c c_{tjs} + \beta_p p_{tj} + \beta_e e_j + \beta_e s_j + \beta_v v_j + \gamma x_{ij} + \varphi_s + \mu_{ij} + \tau_t + \varepsilon_{ijs}, \qquad (1)$$

where subscript t denotes teacher. In equation (1), student test scores are a function the teacher content knowledge, c_{tjs}, teacher pedagogical knowledge, p_{tj}, teacher effort, e_j, teacher skills s_j, a set of observable student, household, and school specific factors x_{ij}, a subject-specific constant φ_s, and three unobserved components – a student-specific component, μ_{ij}, a teacher-specific component, τ_t, and a student by subject specific component, ε_{ijs}.

The most common concern with estimating education production functions of this type is that students with higher aptitude will sort into schools with higher quality teachers and both

---

[20] Regarding the fact that we are computing teacher absence based on a school measure, it is important to note that the school effect is important in predicting whether a given teacher is going to be absent. In fact, the intra-cluster correlation for absence from school or class is equal to 0.32.

will likely sort into better schools (where ε_{ijs} is higher). Estimating the equation by simple OLS would therefore likely lead to biased results because there are unobserved student (and school) characteristics that are correlated both with the teacher quality measures and the left-hand side variable. In addition, if there are other dimensions of teacher quality that matter for student learning that are unobserved but correlated with effort, knowledge and skills, then the estimated coefficients will be biased.

If we want to retain all the teacher quality measures in the estimation, the best we can do is to try to control for all the relevant inputs in the education production function. Since the SDI data contain a wealth of information on students, teachers and schools, we can make some progress here. In particular, we include a student's non-verbal reasoning score, a (admittedly) crude measure of innate ability, whether they had breakfast and their age and gender in the regression. For the teacher we include the same set of characteristics used in the previous section, namely gender, age, experience, and whether they have a university and/or education degree. For the school, we include its location, the pupil-teacher ratio and indices for teaching equipment and infrastructure. In addition, we add country-district fixed effects and day of week and month of survey fixed effects.

If we consider only a subset of the teacher quality measures, namely subject content and pedagogy knowledge, the prospect for identification is improved, because we can now exploit within-student variation. Specifically, we take first differences of equation (1) to arrive at

$$y_{i1} - y_{i2} = \beta_c(c_{t1} - c_{t2}) + \beta_p(p_{t1} - p_{t2}) + \varphi + \Delta\tau_t + \Delta\varepsilon_i \tag{2}$$

where the dependent variable, $y_{i1} - y_{i2}$, is now the difference in student test scores in language and math. Note that content knowledge c_{ts} differs both across and within teachers while pedagogical knowledge is teacher specific. Thus, β_p is identified from the sub-sample of students with different teachers in the two subjects, while β_c is identified off of variation across and within teachers.

This first-differencing approach removes any unobserved student (and school) characteristics that are fixed across subjects. Estimation of equation (2) will generate unbiased estimates of the teacher vector β if (a) the difference in the student-by-subject-specific error terms $(\varepsilon_{i1} - \varepsilon_{i2})$ and (b) the difference in the teacher-specific error term $(\tau_{t1} - \tau_{t2})$ are uncorrelated with the first differenced teacher quality variables. Condition (a) allows for sorting of better students to better teachers, as long as this sorting is not subject-specific. Put simply,

for the specification to be identified, better students may sort into schools with better teachers, but better math students may not sort into schools (or classes) with better math teachers, or at least not disproportionately relative to their sorting in language. We believe that this is a reasonable assumption in the context of primary schools in Sub-Saharan Africa. The first-differencing approach also addresses concerns with matching and tracking of students by performance at the school level, as long as such matching (if it takes place at all) is not based on relative performance in the two subjects. Condition (b) is more restrictive. It states that any unobserved teacher characteristics must be uncorrelated with the observed measures of teacher quality. Otherwise, omitted teacher characteristics such as teacher motivation or effort and skills not captured by our measures could bias the estimates on the observed teacher quality variables.

Given the assumption underlying equation (1), a possible way to overcome such omitted variable bias is to restrict the sample to those that are taught by the same teacher in the two subjects (cf. Metzler and Woessmann, 2012), in which case $\tau_{t1} = \tau_{t2}$. That is, conditional on our empirical model (2) being the true model, by exploiting within student and within teacher variation we can estimate the causal effect of (some components of) teacher quality. Note that restricting the sample comes at a cost as we cannot identify the effects of pedagogical knowledge (unless we are willing to assume that there are no unobservable teacher characteristics correlated with this knowledge and student test scores) and because roughly half of the sample is dropped. Moreover, the remaining schools may be less representative of the schooling environment in the seven countries we have surveyed.

If conditions (a) and (b) hold, or condition (a) holds and condition (b) holds after we restrict attention to students taught by the same teacher in the two subjects, the estimated parameters provide causal estimates of the true effects. But as the explanatory variables are test scores based on tests that are imperfect measures of teachers' true knowledge, the estimates will likely provide lower bounds on the true causal effect. Specifically, if the teacher test scores are measured with classical measurement error, the true effects β_c^* and β_p^* are asymptotically attenuated by the reliability ratio λ (see e.g., Angrist and Krueger, 1999; Metzler and Woessmann, 2012). In general, estimating the reliability ratio λ is problematic. But with the explanatory variables based on test scores, several measures, of which the Cronbach's α (Cronbach 1951) is likely the most widely used, have been developed to estimate reliability. Intuitively, the Cronbach's α is based on the idea that reliability can be estimated by splitting

the test items in two halves and treating them as separate measures of the underlying true concept. The Cronbach's α is the mean of all possible split-half coefficients resulting from different splits of the test and the reliability ratio can be estimated as the ratio of true variance to observed variance of a given test. In Table 7 below, we report both the point estimates from estimating specification (2), $\hat{\beta}$, and the unbiased effects $\hat{\beta}^* = \hat{\beta}\hat{\lambda}$, where $\hat{\lambda}$ is the Cronbach's α estimate of the reliability ratio with first differenced test score data.[21]

The within-student within-teacher results are reported in Table 7. As a benchmark, column (1) reports a simple OLS regression of standardized student test scores (z-scores), on teacher quality and a set of country by district fixed effects as well as survey fixed effects.[22] The point estimates on the four teacher quality dimensions we measure – content knowledge, pedagogical knowledge, pedagogical skills, and effort (absence from class) – are quantitatively large and significant. Students taught by the best possible teachers in the sample, i.e., those who have full scores on the teacher knowledge and pedagogy test and display the full set of skills in the classroom have test scores that are almost 1.5 standard deviations higher than those who are taught by the worst possible teachers. While these results are by no means causal, it should be noted that they exploit only within-district variation in test scores. In column (2) of Table (7), we turn to the estimation of the full model specified in equation (1) that also includes student, teacher and school characteristics. While the coefficients on the teacher variables are reduced somewhat, they are still large and significant. A one standard deviation reduction in absence from class is associated with an increase in student test scores of 0.05 standard deviations. Moving a student from a teacher with no content or pedagogical knowledge and no pedagogical skills to a teacher with full scores on these three dimensions would raise student test scores by almost one standard deviation. Looking at a standard deviation change in knowledge and skills, obviously, yields smaller changes (a one standard deviation change in teacher content knowledge, pedagogical knowledge and skills is associated with roughly a quarter of a standard deviation increase in student test scores), but also implies relatively modest changes in teacher quality.

In column (3), we estimate specification (2). As we exploit within-student variation, we can no longer calculate the effect of classroom absence and pedagogical skills as they are

---

[21] Metzler and Woessmann (2012) derive a reliability ratio for first-differenced test score measures by assuming that the measurement errors are unrelated across the two subjects. We follow their approach here.

[22] Student test scores are standardized in each country. For the teacher pedagogy score, we focus on the lesson preparation score and the evaluating students score, as the comparing writing score is more specific to language teachers only.

absorbed in the student fixed effect. The coefficients are reduced by one-third, suggesting that sorting of students and teachers is indeed relevant, but are still large and significant. Correcting for measurement error doubles the coefficients. Holding teacher effort and classroom skills constant, the estimates imply that moving a student from the bottom 5% of the content and pedagogical distribution of teachers to the top 95%, or essentially moving students from a low quality teacher to one with minimum content and pedagogical knowledge, would result in a 0.7 standard deviation increase in student achievement. This large effect highlights the potentially huge impact of addressing the low teacher quality problem. Specifically, while small changes matter – after all, the point estimates are significant – the scale of the learning problem underlines the importance of identifying interventions and reforms with potential to substantially raise student learning outcomes. The findings in column (3) suggest that interventions that would result in students being taught by teachers with at least minimum subject content and pedagogical knowledge are such interventions.

A remaining concern with the findings reported in Column (3) is that unobserved teacher characteristics might bias the estimates. Column (4) reports the results of the within-student within-teacher fixed effects specification. Column (5) goes one step further by restricting the sample to students taught by the same teacher in both subjects and where there is only one grade 4 classroom, which effectively rules out any bias from sorting between classrooms (subject-specific or otherwise) within the school. As evident, the point estimates on teacher content knowledge remain largely unchanged.

9. How do private schools compare to public schools?

In many low-income countries—including many in Sub-Saharan Africa—private schools are a growing part of the education landscape. Private schools respond to a demand on the part of parents, either because parents can't access public schools or because the public schools do not provide services of acceptable quality in their judgement. The internal incentives within private schools (e.g. teachers vis a vis principals or school managers), and the accountability of private schools with respect to parents, have been argued as reasons for why performance may be higher in these schools. In principle, private school performance need not be uniformly high, though: some markets may support relatively low-quality private schools (Andrabi, Das, and Khwaja 2015). In addition, however, in many settings, the growth in enrollment in private schools may have diluted the quality of public education, resulting in the exit of better-off

children and increased economic stratification (Hsieh and Urquiola 2006). This, in turn, may reduce pressure on public schools to supply quality education (Fiske and Ladd 2003).

While the SDI data cannot be used to assess the validity of these various arguments, we can compare the performance of public and private schools in the sample. In general, indicators in private schools tend to be better: Private school teachers tend to put in more effort, show more knowledge, and exhibit better teaching practices than their public counterparts (Table 8). At the same time, it is important to note that private schools are not able to overcome many of the poor service delivery issues faced by public schools—the issues appear to be systemic. Indeed, even in the private sector, one-third of teachers are absent from the classroom. While the taught school day is four hours long on average, i.e., more than one hour longer than in the average public school, it is still well short of scheduled time. In addition, while teachers in private schools have significantly higher test scores, their pedagogical knowledge is similar to their public school counterparts.

The better performance of private school teachers is reflected in their students' learning. The student score in mathematics and language is one-third and two-thirds higher in private schools. A student in private school is 50 percent more likely to be able to read a word, and his or her reading comprehension score is three times as high as those of public school students. However, there is also some evidence that at least some of this superior performance is due to sorting of students: the non-verbal reasoning score in private schools is 13 percent higher than in public schools.

10. Discussion and Conclusion

In this paper, we report on what primary school teachers in Africa know and do, using representative data from an ongoing survey program: The Service Delivery Indicators data. The findings provide a concerning picture of teacher effort, knowledge, and skill, although there is significant variation in teacher quality, both within and across countries. Even taking this variation into account, however, if "adequate" teaching is defined by students that during most of their scheduled time are being taught by teachers with at least basic pedagogical knowledge and minimum subject knowledge in language and mathematics, then very few public primary school students, in the seven countries we surveyed, benefit from decent quality education.

We argue, and partly show, that this lack of quality education in turn can help explain why so many students learn little and complete their primary education lacking even basic reading, writing, and arithmetic skills. Potential human capital for cohorts of students is consequently lost.

At a general level, given the results presented here, it is easy to list what governments "should" do to improve service performance in the education sector, assuming they have the capacity and commitment to do so. Specifically, the results highlight the importance of attracting talented candidates to teaching and preparing them to teach the curriculum effectively. The results also highlight the need to put in place effective schemes that ensure high effort and continued upgrading of knowledge and skills. Innovative programs along these lines may not only train teachers, but also adjust curriculum materials to better suit student needs (USAID 2016). As of yet, unfortunately, we have little systematic evidence of how cost-effective reforms of this type in practice should be implemented in order to be effective at scale.

More specifically, these findings highlight the need to understand and assess complementarities between teacher effort, ability, and skills in generating high quality education. We show that teachers, on average, both teach too little and lack the necessary skills and knowledge to teach effectively when they actually teach. It is difficult to think of any intervention in education that would have dramatic effects on learning outcomes if it does not simultaneously address low teacher effort, low knowledge, and poor skills. For example, there is by now strong evidence that both teacher effort and skills can be raised, leading to improved learning outcomes (see, for example, the reviews in Kremer, Brannen, and Glennerster, 2013; and Glewwe and Muralidharan, 2015). But neither of these alone will likely be enough to significantly change the quality of education when many teachers do not even master their students' curriculum.

The results therefore imply that there is a need to address issues at the level of a system as a whole. The problems identified are large in magnitude and broad in scope, suggesting that specific interventions, even if deployed in a large number of schools, are unlikely to make much of a difference at scale. Teacher recruitment, preparation, deployment, incentives and motivation, along with ongoing professional development, will all likely matter for creating a cadre of professional educators who provide high quality education. The key research and operational challenge is to find approaches that deliver the combination of these that works for a country's particular context—and ultimately deliver learning.

References

ASER. 2013. Annual Status of Education Report (Rural) 2013. ASER Center. New Delhi.

Andrabi, T., J. Das, and A. I. Khwaja, (2015), Delivering Education: A Pragmatic Framework for Improving Education in Low-Income Countries, Policy Research Working Paper 7277, The World Bank.

Banerjee, A., R. Banerji, J. Berry, E. Duflo, H. Kannan, S. Mukerji, M. Shotland, and M. Walton, 2016, Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of "Teaching at the Right Level" in India, Working Paper, MIT.

Bau, N., and J. Das, (2016), The Misallocation of Pay and Productivity in the Public Sector: Evidence from the Labor Market for Teachers, Working Paper, The World Bank.

Baum, Donald Rey, Laura Lewis, Oni Lusk-Stover, Harry Anthony Patrinos. 2014. "What matters most for engaging the private sector in education: a framework paper." Systems Approach for Better Education Results (SABER) Working Paper Series No. 8. Washington, DC : World Bank Group.

Bold, T., M. Kimenyi, G. Mwabu, A. Ng'ang'a and J. Sandefur, (2016), Interventions and Institutions: Experimental Evidence on Scaling Up Education Reforms in Kenya, Working Paper, IIES, Stockholm University.

Bruns, B., Filmer, D., and Patrinos, H. A. (2011). Making Schools Work New Evidence on Accountability Reforms. Number 2270 in World Bank Publications. The World Bank.

Bruns, Barbara, and Javier Luque. (2014). Great Teachers: How to Raise Student Learning in Latin America and the Caribbean. Washington, DC: World Bank.

Bruns, B., Mingat, A., and Rakotomalala, R. (2003). Achieving Universal Primary Education by 2015 : A Chance for Every Child. Number 15121 in World Bank Publications. The World Bank.

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., and Rogers, F. H. (2006). Missing in action: teacher and health worker absence in developing countries. *The Journal of Economic Perspectives*, 20(1):91–116.

Chetty, R., J.N. Friedman, J.E. Rockoff, (2014), Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood, *American Economic Review*, 104 (9): 2593-2632.

Coe, Robert, Aloisi, Cesare, Higgins, S. and Lee Eliot Major (2014), What makes great teaching? Review of the underpinning research, Sutton Trust.

Duflo, E., P. Dupas, and M. Kremer, (2015), School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools, *Journal of Public Economics*, 123: 92–110

Evans, D., and A. Popova, (2016), What Really Works to Improve Learning in Developing Countries: An Analysis of Divergent Findings in Systematic Reviews, World Bank Research Observer, first published online 5th of October 2016.

Fiske, Edward B. and Helen F. Ladd, (2003), Balancing Public and Private Resources for Basic Education: School Fees in Post-Apartheid South Africa, Working Paper Series SAN03-03, Terry Stanford Institute of Public Policy. Duke University.

Glewwe, P., and K. Muralidharan, (2015), Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications, RISE-WP-15/001.

Hanushek, E.A. and L. Woessmann, (2015), *The Knowledge Capital of Nations: Education and the Economics of Growth*, MIT Press.

Hsieh, Chang-Tai and Miguel Urquiola (2006), The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program, *Journal of Public Economics*, 90(8-9):1477-1503.

Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolata, S., Cappelle, F.V. (2010). SACMEQIII Project Result: Pupil Achievement Levels in Reading and Mathematics. Working Document Number 1. Paris: SACMEQ.

Johnson, David, Cunningham, Andrew and Rachel Dowling (2012), Teaching Standards and Curriculum Review, World Bank

Ko, James and Sammons Pamela (2013), Effective teaching: a review of research and evidence, CfBT Education Trust

Kremer, M., C. Brannen and R. Glennerster, (2013), The Challenge of Education and Learning in the Developing World, *Science* 340:297-300.

Krishnaratne, S., H. White, and E. Carpenter, (2013), Quality education for all children? What works in education in developing countries, International Initiative for Impact Evaluation (3ie) Working Paper 20. New Delhi.

Malpel, Jaques. 2016. *PASEC 2014: Education system Performance in Francophone Sub-Saharan Africa*. Programme d'Analyse des Systèmes Educatifs de la CONFEMEN. Dakar, Sénégal.

Murnane, R., and A. Ganimian, (2014), Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Evaluations". NBER Working Paper no. 20284.

Daniel Muijs, Leonidas Kyriakides, Greetje van der Werf, Bert Creemers, Helen Timperley & Lorna Earl (2014) State of the art – teacher effectiveness and professional learning, *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 25:2, 231-256

Ozier, O. (2017) Measuring non-verbal reasoning across sub-Saharan Africa. (mimeo)

Stallings, J. (1980). Allocated Academic Learning Time Revisited, or Beyond Time on Task. *Educational Researcher*, 9:11–16.

UIS. 2011. *Financing Education in Sub-Saharan Africa: Meeting the Challenges of Expansion, Equity and Quality.* UNESCO/UIS. Montreal.

UIS/ISU. 2013.

UNESCO. 2013. *The Global Learning Crisis: Why Every Child Deserves a Quality Education.* UNESCO. Paris.

USAID. 2016. *Tusome Early Grade Reading Activity*. Fact Sheet.
http://usaid.gov/sites/default/files/documents/1860/Tusome%20%20Factsheet%202016.pdf

S. Vieluf, D. Kaplan, E. Klieme and S. Bayer (2012), "Teaching Practices and Pedagogical Innovations: Evidence from Tali", OECD Publishing.
http://dx.doi.org/10.1787/9789264123540-en

Table 1: Student Knowledge

	All	Kenya	Mozambique	Nigeria	Senegal	Tanzania I	Tanzania II	Togo	Uganda
Literacy									
Pupil can read a letter (%)	62	89	38	50	79	46	51	72	72
Pupil can read a word (%)	53	82	21	21	74	54	55	57	58
Pupil can identify words (%)	30	50	21	44	18	6	10	57	30
Pupil can read a sentence (%)	29	71	13	19	51	12	11	16	35
Pupil can read paragraph (%)	11	26	7	6	25	7	2	9	6
Comprehension, score out of 100	15	40	5	10	16	5	13	10	21
Numeracy									
Pupil can recognize numbers (%)	88	97	81	61	97	90	92	96	91
Pupil can order numbers (%)	45	71	20	20	66	43	44	49	46
Pupil can add single digits (%)	76	92	48	51	86	94	79	74	81
Pupil can add double digits (%)	56	82	18	30	77	70	60	61	53
Pupil can add triple digits (%)	54	85	8	20	78	64	59	61	53
Pupil can subtract single digits (%)	66	87	28	45	79	81	73	60	74
Pupil can subtract double digits (%)	30	59	5	18	39	40	38	16	24
Pupil can multiply single digits (%)	26	48	4	18	31	39	37	10	21
Pupil can multiply double digits (%)	10	5	0	3	39	13	12	3	1
Pupil can multiply triple digits (%)	6	1	0	2	19	9	9	3	1
Pupil can divide single digits (%)	33	57	9	18	42	39	38	30	34
Pupil can divide double digits (%)	15	33	3	9	20	16	18	7	11
Pupil understands division (%)	15	27	11	9	8	13	20	19	13
Pupil can solve math story (%)	5	10	0	3	5	8	9	6	2
Pupil can complete a sequence (%)	10	26	4	11	3	1	15	8	10
No. of students	19561	2368	1731	3723	1440	1786	3909	1423	3181

Notes: The table reports the share of students in government primary schools in each country who have certain competencies. Note that in Tanzania II, the results in language are for 2736 students as the remainder were tested in Swahili. All individual country statistics are calculated using country-specific sampling weights. The average for all countries is taken by averaging over the country columns. Hence, each country is given equal weight. Further details on the construction of the variables and sampling weights are available in an Appendix available from the authors upon request.

Table 2: Teacher Absence

	All	Kenya	Mozambique	Nigeria	Senegal	Tanzania I	Tanzania II	Togo	Uganda
Absence from class (%)	44	48	56	23	31	53	47	40	57
Absence from school (%)	23	15	45	17	16	23	15	23	28
No. of teachers	16543	2311	991	2968	1222	1740	3518	776	3017
Scheduled teaching time	5h 27m	5h 36m	4h 21m	4h 43m	4h 36m	5h 47m	5h 55m	5h 24 m	7h 13m
Time spent teaching	2h 49m	2h 31m	1h 43m	3h 10m	3h 5m	2h 32m	3h 16m	3h 9m	3h 2m
No. of schools	2001	238	200	435	145	176	344	144	319
Orphaned classrooms (%)	33	38	31	26	-	-	36	24	45
No. of schools	1647	234	150	430	-	-	392	127	314

Notes: The table reports the absence rate for all teachers, the scheduled teaching time, actual teaching time and number of orphaned classrooms for all government schools. Teachers are marked as absent from school if during the second unannounced visit, they are not found anywhere on the school premises. Otherwise, they are marked as present. Teachers are marked as absent from class if during the second unannounced visit, they are absent from school or present at school but absent from the classroom. Otherwise, they are marked as present. The scheduled teaching time is the length of the school day minus break time. Time spent teaching adjusts the length of the school day by the share of teachers who are present in the classroom, on average, and the time the teacher spends teaching while in the classroom. The orphaned classrooms measure is the ratio of the classrooms with students but no teacher to the number of classrooms with students with or without a teacher. All individual country statistics are calculated using country-specific sampling weights. The average for all countries is taken by averaging over the country columns. Hence, each country is given equal weight. Further details on the construction of the variables and sampling weights are available in an Appendix available from the authors upon request.

Table 3: Teachers reaching "minimum thresholds" on knowledge assessment

Percentage (%) of teachers with …	All	Kenya	Mozambique	Nigeria	Senegal	Tanzania I	Tanzania II	Togo	Uganda
Subject knowledge: Language									
80% of knowledge equivalent to a 4th grader	66	94	73	26	81	60	49	57	90
Minimum knowledge	7	34	0	0	6	0	1	0	16
No. of teachers	3741	633	315	681	188	97	697	604	555
Subject knowledge: Maths									
Minimum knowledge	68	93	50	51	75	78	79	49	70
No. of teachers	3922	632	312	691	188	91	858	604	581
Pedagogy knowledge									
Minimum Knowledge General Pedagogy	9	10	3	1	-	-	36	3	4
Minimum Knowledge Assessing Students	0	0	0	0			0	0	0
No. of teachers	4751	893	324	769			1264	604	945

Notes: The table reports minimum knowledge indicators for teachers in grade 4 or who taught grade 3 in the previous year in government schools. A language teacher is defined as having 80% of knowledge equivalent to a fourth grader in language if he/she score at least 80% on the tasks (of similar difficulty) that were contained on both the student and the teacher tests. A language teacher is defined as minimum knowledge if he/she score at least 80% on the grammar, Cloze test and correcting a student's composition task of the language assessment. A mathematics teacher is defined as having minimum knowledge (=80% of knowledge equivalent to a 4th grader) if he/she score at least 80% on the tasks (of similar difficulty) that were contained on both the student and the teacher test and (roughly) similar in difficulty. In language and mathematics, a teacher is denoted as having minimum knowledge if they have 80% of knowledge required for teaching a fourth grader. A teacher in any subject is defined as having minimum knowledge of general pedagogy if they score least 80% on the tasks that relate to general pedagogy (factual text comprehension and being able to formulate learning outcomes and lesson aims). A teacher in any subject is defined as having minimum knowledge assessing students if they score least 80% on the tasks that relate to assessment (comparing students' writing and monitoring progress among a group of students). All individual country statistics are calculated using country-specific sampling weights. The average for all countries is taken by averaging over the country columns in the table. Hence, each country is given equal weight. Further details on the construction of the variables and sampling weights are available in an Appendix available from the authors upon request.

Table 4: Teacher's performance on specific item groups of knowledge

	All	Min	Max
Language			
Spelling task, score out of 100	86	86 (Ta)	86 (Ta)
Grammar task, score out of 100	79	58 (Ni)	92 (Ke)
Cloze task, score out of 100	44	27 (To)	66 (Ke)
Correct composition task, score out of 100	26	9 (Mo)	50 (Ke)
No. of teachers	3770		
Math			
Teacher can add double digits (%)	91	82 (Mo)	98 (Ke)
Teacher can subtract double digits (%)	77	59 (Ni)	92 (Se)
Teacher can multiply double digits (%)	68	44 (Mo)	89 (Se)
Teacher can solve simple math story problem (%)	55	17 (Mo)	91 (Se)
Teacher understands a Venn diagram (%)	41	19 (To)	70 (Ke)
Teacher can interpret data in a graph (%)	25	12 (To)	62 (Ke)
Teacher can solve algebra (%)	35	3 (Mo)	74 (Ke)
Teacher can solve difficult math story problem (%)	15	7 (Se)	22 (Ta)
No. of teachers	3957		
Pedagogy			
Factual text comprehension, score out of 100	46	23 (Mo)	78 (Ta)
Preparing a lesson plan, score out of 100	31	15 (Ni)	58 (Ta)
Design Lesson:			
aims and learning outcomes, score out of 100	22	11 (Ni)	41 (Ta)
formulate question that checks understanding, score out of 100	17	5 (Ni)	52 (Ta)
formulate question that applies learned to other contexts, score out of 100	7	3 (Ni)	15 (Ta)
Assessing students' abilities	18	8 (Ni)	32 (Ke)
Evaluating students' progress	12	5 (Ni)	26 (Ke)
No. of teachers	5181		

Notes: The table presents scores on specific tasks for teachers in government schools in grade 4 or who have previously taught grade 3. The scores on the language test are computed for teachers teaching language, the scores on the mathematics test are computed for teachers teaching mathematics and the scores on the pedagogy section are computed for teachers teaching either subject. The initials of the countries with the lowest and highest score for each item are given in brackets. All individual country statistics are calculated using country-specific sampling weights. The average for all countries is taken by averaging over the country columns in the table. Hence, each country is given equal weight. Further details on the construction of the variables and sampling weights are available in an Appendix available from the authors upon request.

Table 5: Teacher Skills and Practices in the Classroom

	All	Kenya	Mozambique	Nigeria	Tanzania	Togo	Uganda
Teacher introduces and summarizes the topic of the lesson (%)	41	62	16	55	46	36	30
Lesson appears planned to enumerator (%)	64	75	71	60	67	74	37
Teacher asks a mix of lower and higher order questions (%)	31	31	14	36	32	30	44
Teacher gives positive feedback, praise, corrects mistakes (%)	52	70	32	43	59	35	75
Teacher engages in all of the above practices (%)	9	17	1	10	12	3	5
No. of classrooms	1558	181	197	428	338	140	281

Notes: The table presents teacher practices in the classroom in government schools in grade 4. The information is not available for Senegal and Tanzania (1st survey). 'Teacher introduces and summarizes the topic of the lesson' is a dummy set to 1 if the teacher introduces and summarizes the lesson and zero otherwise. 'Lesson appears planned to enumerator' is a dummy set to 1 if the lesson appears planned to the enumerator and zero otherwise. 'Teacher asks a mix of lower and higher order questions' is a dummy set to 1 if the teacher asked questions that required learners to recall information and the teacher asked learners to carry out tasks which allowed them to demonstrate their understanding of what they had learned during the lesson and the teacher asked questions that required learners to apply information to new topics and the teacher asked questions which required learners to use their creativity and imagination, and zero otherwise. 'Teacher gives positive feedback and praise and corrects mistakes' is a dummy set to 1 if the teacher gave feedback of praise at least once and corrected a mistake at least once and did not scold at a mistake more than once. All individual country statistics are calculated using country-specific sampling weights. The average for all countries is taken by averaging over the country columns in the table. Hence, each country is given equal weight. Further details on the construction of the variables and sampling weights are available in an Appendix available from the authors upon request.

Table 6: Who are the good teachers?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Effort (absence)		Content knowledge		Pedagogical knowledge		Pedagogical skills	
Female	-0.08***	-0.06***	-0.01	-0.005	0.00	0.00	0.08***	0.07***
	(.01)	(.01)	(.01)	(.005)	(.01)	(.01)	(.01)	(.01)
Born in district	0.03***	0.02**	-	-	-	-	-	-
	(.01)	(.01)						
Experience	-0.00	0.002**	0.00	-0.00	-0.001***	0.00	-0.001***	0.00
	(.01)	(.001)	(.01)	(.001)	(.00)	(.00)	(.00)	(.001)
Age	-0.00	0.00	0.00	0.00	-0.001***	-0.001***	0.002***	0.002
	(0.01)	(.01)	(0.01)	(.01)	(0.00)	(.00)	(0.00)	(.001)
Degree	0.03***	0.03**	0.04***	0.04**	0.04***	0.04***	0.07***	0.05***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.02)	(.02)
Contract teacher	-0.07***	-0.03**	0.005	0.01*	-0.01	-0.01	-0.01	0.01
	(.02)	(.02)	(.007)	(.01)	(.01)	(.01)	(.02)	(.03)
Head teacher	0.19***	0.18***	-	-	-	-	-	-
	(.01)	(.01)						
Specification	Binary	Multivariate	Binary	Multivariate	Binary	Multivariate	Binary	Multivariate
Schools		2,052		2,027		1,706		1,625
Observations		15,688		8,228		7,611		1,625

Notes: The table presents estimates from binary (columns 1, 3, 5, 7) and multivariate (columns 2, 4, 6, 8) regressions. Each row in columns 1, 3, 5, 7 is based on a separate regression with the dependent variable (reported at the top of the column) regressed on the explanatory row variable. The information on all teacher quality variables are not available for Senegal and Tanzania (1st survey) and data for those samples are therefore not included. 'Effort (absence from classroom)' is a dummy set to 1 if the teacher is not found in the classroom and zero otherwise. 'Content knowledge' measure the percentage of correct questions on the teacher content knowledge test (language for language teachers, math for math teachers, average of language and math for teachers teaching both subjects), ranging from 0 to 1. 'Pedagogical knowledge' is average score on the pedagogy part of the test, ranging from 0 to 1. 'Pedagogical skills' is the average of the four skills presented in Table 5, i.e. structuring, planning, asking lower and higher order questions and giving feedback, ranging from zero to 1. Robust standard errors clustered by school in parentheses. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level. The variable born in district was not collected for teachers tested for content and pedagogical knowledge and skills.

Table 7: Student learning and teacher quality

	(1)	(2)	(3)	(4)	(5)
Regression estimates:					
Teacher content knowledge	0.54***	0.37***	0.25***	0.20***	0.23***
	(.06)	(.06)	(.05)	(.06)	(.07)
Teacher pedagogical knowledge	0.50***	0.35***	0.23**		
	(.10)	(.09)	(.11)		
Teacher pedagogical skills	0.38***	0.25***			
	(.05)	(.05)			
Teacher absence from class	-0.14***	-0.08**			
	(.04)	(.03)			
Language	-0.03*	-0.04***	-0.03***	-0.01	-0.02
	(.02)	(0.01)	(.01)	(.02)	(.02)
Measurement-error corrected estimates:					
Teacher content knowledge			0.60	0.64	0.64
Teacher pedagogical knowledge			0.32	-	-
Student controls	No	Yes	-	-	-
School controls	No	Yes	-	-	-
Teacher controls	No	Yes	Yes	-	-
Country by district FE	Yes	Yes	-	-	-
Student FE	No	No	Yes	Yes	Yes
Survey fixed effects	Yes	Yes	-	-	-
Sample	All	All	All	Same teacher	Same teacher 1 class
Schools	1,688	1,670	1,950	904	760
Observations	25,777	25,424	29,597	15,128	12,748

Notes: Point estimates from regressing student learning (standardized test scores) on dimensions of teacher quality. 'Teacher content knowledge' measures the percentage of correct questions on the teacher content knowledge test (language for language teachers, math for math teachers), ranging from 0 to 1. 'Teacher pedagogical knowledge' is average score on the pedagogy part of the test, ranging from 0 to 1. 'Teacher pedagogical skills' is the average of the four skills presented in Table 5, i.e. structuring, planning, asking lower and higher order questions and giving feedback, ranging from zero to 1. 'Teacher absence from class' is the share of teachers in a school absent from class based on an unannounced survey. Student controls are age, gender, and whether the student had breakfast the day of the test. School controls are whether the school is situated in an urban area, index of available school equipment (students have access to pens and exercise books and functioning blackboard in the classroom), index of available school infrastructure (toilets accessible, clean and private and visibility in the classroom) and whether the school is private. Teacher controls include age, gender, experience, and whether the teacher has an education college or university degree. Robust standard errors clustered by school in parentheses. *Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

Table 8: Comparing public and private schools

	Public	Private
Teachers		
Absence from class (%)	42	27
Time Spent Teaching	3h 18m	4h 2m
Score Language	50	55
Score Mathematics	50	58
Score Pedagogy	22	25
Good practice in the classroom	9	12
Students		
Score Language	44	73
Score Maths	42	56
Score NVR	53	61

Notes: The table presents measures of teacher quality and student learning and characteristics in public and private schools. For the public schools, the average is taken over those samples (countries), for which a sizeable private school sample is available, namely Kenya, Nigeria, Togo and Uganda. Teacher variables: 'Absence from class' the average share of teachers in a school absent from class based on an unannounced survey; 'Time spent teaching' is the average time teachers spend teaching in the classroom; 'Score Language' is average score on the language test (0-100); 'Score Mathematics' is the average score on the mathematics test (0-100); 'Score Pedagogy' is the average score on the pedagogy part of the test (0-100); Good practice in the classroom is the average share of teachers that practice all four of the skills in table 5.. Student variables: 'Score Language' is the average score on the language part of the student test (0-100); 'Score Maths' is the average score on the mathematics part of the student test (0-100); 'Score NVR' is the average score on the non-verbal reasoning part of the test (0-100).