Teacher Subject Knowledge, Didactic Skills, and Student Learning in Francophone Sub-Saharan Africa*

Jan Bietenbeck[†], Natalie Irmert[‡], Mohammad H. Sepahvand[§]

This version: May 30, 2022

Abstract

We study the effects of two dimensions of teacher quality, subject knowledge and didactic skills, on student learning in francophone Sub-Saharan Africa. We use data from an international large-scale assessment in 14 countries that include individual-level information on student achievement and country-level measures of teacher subject knowledge and didactic skills in reading and math. Exploiting variation between subjects in a student fixed-effects model, we find that teacher subject knowledge has a large positive effect on student achievement, whereas the effect of didactic skills is comparatively small and not statistically significant at conventional levels. Together, the two dimensions of teacher quality account for 36 percent of the variation in average student achievement across countries.

Keywords: international learning gaps, teacher quality, Sub-Saharan Africa
JEL codes: I21, I25, O15

*We thank Simon Wiederhold and the audience at the spring 2022 Copenhagen Education Network Workshop for helpful comments. Jan Bietenbeck is grateful for funding from the Jan Wallanders och Tom Hedelius stiftelse samt Tore Browaldhs stiftelse. Mohammad H. Sepahvand is grateful for funding from the Swedish Research Council.

†Lund University and CESifo, DIW Berlin, IZA.

‡Lund University.

§Lund University and European Center for Advanced Research in Economics and Statistics (ECARES), Université libre de Bruxelles.

1 Introduction

A growing literature in economics shows that differences in human capital account for a large part of cross-country differences in economic performance (e.g. Jones 2014, Hendricks & Schoellman 2018). Especially cognitive skills, as measured by student performance on international standardized tests, are a crucial driver of economic growth (Hanushek & Woessmann 2012*a*,*b*, Hanushek 2013). In Sub-Saharan Africa, which includes the majority of the least developed countries, there has been a dramatic rise in school enrollment over the past two decades. However, standardized tests reveal that children in this region are learning very little in school, which limits the positive effect this rapid educational expansion has on growth (World Bank 2018). Importantly, this low general level of learning masks considerable heterogeneity across countries: for example, whereas 46 percent of sixth-grade students in Niger have difficulties reading a simple sentence, this figure stands at 12 percent in neighboring Burkina Faso (PASEC 2020). Understanding the causes of these international differences is important for economic and education policy, but so far only very little research has attempted to identify the causal factors behind learning gaps between Sub-Saharan African countries.

In this paper, we study the role of one potential factor behind these gaps: teacher quality. Teachers are widely seen as the most important school-based input into learning (e.g. Hanushek & Rivkin 2006, 2012), but comparable international measures of teacher quality are rare. We use novel data from the Program for the Analysis of Education Systems (*Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN*, PASEC), which conducts large-scale learning assessments in francophone countries in Sub-Saharan Africa. In 2019, PASEC assessed the reading and math skills of nationally representative samples of sixth-grade students in 14 countries. Unusually, it also assessed the subject knowledge and didactic skills of their teachers. Previous research has shown that subject knowledge, which refers to teachers' mastery of the knowledge that they are expected to teach, affects student learning within both developed and developing countries (e.g. Rockoff et al. 2011, Metzler & Woessmann 2012, Bietenbeck et al. 2018). Correlational evidence also links didactic skills, which describe teachers' ability to adapt subject knowledge for teaching purposes, to student achievement (e.g. Hill et al. 2005, Baumert et al. 2010, Sadler et al. 2013). In our analysis, we ask whether differences between countries in these two dimensions of teacher quality contribute to the large international learning gaps in Sub-Saharan Africa.

The identification of causal determinants of cross-country differences in learning is complicated because countries and their education systems differ in numerous dimensions, many of which are unobserved. To overcome this challenge, we exploit the fact that PASEC assessed students and teachers in two subjects, reading and math. For each subject, the data let us observe student achievement at the individual level and average teacher subject knowledge and didactic skills at the country level. Our main regressions at the student level relate the *difference* in student achievement between reading and math to the corresponding *differences* in teacher skills. This is equivalent to introducing student fixed effects, and it implies that we control for all potential student-, school-, and country-level confounders that do not vary between the two subjects. Our regressions also control for subject-specific factors that could still bias these estimates, such as numeracy and literacy in the general population.

We find that cross-country differences in teacher quality predict international learning gaps.

We first estimate the effects of teacher subject knowledge and teacher didactic skills in two separate regressions. In these specifications, a one standard deviation (SD) increase in subject knowledge is estimated to raise student achievement by 0.71 SD, and a one SD increase in didactic skills is estimated to raise student achievement by 0.58 SD. However, when we include both dimensions of teacher quality in the same regression, only the effect of subject knowledge prevails: in this horse race specification, a one SD increase in subject knowledge raises student achievement by 0.69 SD, whereas the effect of didactic skills is comparatively small at 0.07 SD and not statistically significant at conventional levels. Together, the two dimensions of teacher quality account for 36 percent of the variation in average student achievement across countries.

Our paper contributes to a growing body of research on causal determinants of international learning gaps. This literature has found that differences in school autonomy (Hanushek et al. 2013), instruction time (Lavy 2015, Bietenbeck & Collins 2020), student testing (Bergbauer et al. 2021), and time and risk-taking preferences (Hanushek et al. 2021) explain part of these gaps.[1] Moreover, in work that is closely related to our paper, Hanushek et al. (2019) show that teacher cognitive skills predict cross-country differences in student achievement. A common feature of all of these papers is their focus on middle- and high-income countries outside of Sub-Saharan Africa. In contrast, Bietenbeck et al. (2018) and Bold et al. (2019) pool data on teacher-level subject knowledge linked to student-level achievement from several Sub-Saharan African countries. While they show that subject knowledge affects achievement within these pooled samples, they do not estimate whether differences in teacher skills explain differences in learning *between* countries. We contribute to this literature by providing the first evidence on causal determinants of international learning gaps in low-income countries. Moreover, we provide some of the first causal evidence on the effect of didactic skills.[2]

2 Background

Francophone Sub-Saharan African countries are among the least-developed countries in the world, with more than a quarter of the population living below the international poverty line of 1.90 USD per day (World Bank 2021*c*).[3] The problems associated with this widespread poverty are manifold and include poor nutrition and health, child labor, institutional instability, and violent conflict. While these challenges are common to all countries in the region, there are important international differences: for example, while more than 40 percent of the population in Niger lives in poverty, only three percent of the population in Gabon does (World Bank 2021*c*). Similarly, GDP per capita ranges from 239 USD in Burundi to 6,882 USD in Gabon (World Bank 2021*a*). Countries also experience very different levels of violence and social unrest, with

[1]Many other studies use data from international student assessments but do not explicitly try to explain cross-country differences in learning. For an overview of the literature on international differences in student achievement, see Woessmann (2016).

[2]The effect of didactic skills on student achievement has received relatively little attention by economists. Within educational science, a few studies have found positive associations between didactic skills and student achievement (e.g. Hill et al. 2005, Marshall et al. 2009, Baumert et al. 2010, Ngo 2013, Sadler et al. 2013, Cueto et al. 2017). However, these associations are unlikely to capture causal effects because the underlying analyses do not account for the likely sorting of students between and within schools.

[3]The statistics in this Section refer to the fourteen francophone Sub-Saharan African countries covered by the PASEC data, which are listed in Section 3.

the Sahel region being especially affected.

Against this background, improvements in education are often seen as a key means to boost economic and social development in francophone Sub-Saharan Africa. In the past two decades, the region has made substantial progress in increasing educational attainment: many countries eliminated fees for primary schooling and partly as a consequence, gross enrollment rates in primary education in most countries rose above 100 percent (World Bank 2021*d*). Notwithstanding this success, these figures hide that the *quality* of schooling is often poor: indeed, many students complete their six-year primary education without having acquired basic literacy and numeracy skills (World Bank 2018). As noted in the introduction, there also are large differences in average levels of learning in school between countries (PASEC 2020).

There are several potential explanations for these low general levels of learning in francophone Sub-Saharan Africa and for the differences in learning levels between countries. One possibility is that physical school resources are inadequate: for example, textbooks are often not available for all students, and many schools do not have electricity. However, a large body of research has shown that such physical resources play only a limited role in explaining student achievement (e.g. Hanushek & Woessmann 2011). Another possibility is that a lack of qualified teachers hampers learning in schools. Indeed, a growing school-age population and increasing enrollment rates have dramatically increased the demand for teachers. But in many countries, qualified teachers are scarce and vacant positions can often only be filled with unqualified candidates (World Bank 2018). In our analysis below, we therefore investigate whether cross-country differences in two important dimensions of teacher quality, subject knowledge and didactic skills, explain international gaps in student learning.

3 Data

3.1 The PASEC assessments

The Conference of Ministers of Education of French-Speaking Countries (*Conférence des Ministres de l'Éducation des États et Gouvernement de la Francophonie*, CONFEMEN) created PASEC in 1991 with the purpose of conducting regular assessments of student skills in its member countries. The program initially focused on country-specific assessments which were not internationally comparable, but it shifted to a standardized international format similar to that of the OECD's Programme for International Student Assessment (PISA) for its two most recent assessments in 2014 and 2019. In those two years, PASEC assessed nationally representative samples of sixth-grade students on the reading and math skills that they should have acquired by the end of primary school. Only in 2019, it also tested their teachers on the same end-of-primary-school skills as well as on their didactic skills. Moreover, it assessed additional smaller samples of second-grade students on lower-level reading and math skills. In this paper, we focus on the samples of sixth-grade students tested in 2019 because of the immediate relevance of the observed measures of teacher quality for these students' test performance.

PASEC 2019 used a three-stage sampling design to draw nationally representative samples of sixth-grade students in the following 14 countries: Benin, Burkina Faso, Burundi, Cameroon, Chad, Republic of the Congo, Democratic Republic of the Congo, Gabon, Guinea, Ivory Coast, Madagascar, Niger, Senegal, and Togo.[4] In the first stage, primary schools in each country were selected with a probability proportional to enrollment. In the second stage, one class was chosen at random from all sixth-grade classes in selected schools. In the third stage, 25 students from each class were randomly selected to participate in the assessment. Moreover, all teachers employed at the primary schools selected in the first stage were assessed on their subject knowledge and didactic skills. In our empirical analysis below, we always use appropriate sampling weights in order to account for this complex sampling design.

Students participating in PASEC 2019 were assessed on their reading and math skills using standardized multiple-choice tests, which covered core competencies that students should have acquired by the end of primary school. In particular, the reading tests assessed students in the following two areas: (1) understanding isolated words and sentences and (2) text comprehension. The math tests assessed students in the following three areas: (1) arithmetic, (2) measurement, and (3) geometry and space. The language of assessment was French, with a few exceptions where the tests were translated into the local language of instruction. As is commonly the case for other international assessments, PASEC used item response theory to place student test scores on a common international scale. This scale was first introduced in PASEC 2014 and was normalized to have mean 500 and SD 100 across the countries participating in that wave. To ensure comparability over time, the scores from PASEC 2019 were put onto this same scale.

PASEC 2019 assessed primary school teachers' subject knowledge in reading and math using multiple-choice tests. The assessment evaluated teachers' mastery of the skills that are expected from students at the end of primary school and covered areas that largely overlapped with those in the sixth-grade student tests. Thus, teacher performance on the assessment reflects subject knowledge that is likely highly relevant for student learning in primary school. Like with the student tests, teachers' scores on the subject knowledge test were placed on a common international scale with mean 500 and SD 100.

The assessment of teachers' didactic skills was based on Shulman's model of pedagogical reasoning (Shulman 1986, 1987). The model defines a teacher's quality as her ability to draw a link between pure subject knowledge and pedagogical competencies, and hence to adapt subject knowledge for teaching. Shulman (1986) derives five didactic skills that are required for this process, which PASEC pooled into the following two dimensions: (1) planning a lesson for pre-specified learning objectives and (2) identifying the types and sources of students' errors. The assessment evaluated teachers' subject-specific skills on these two dimensions separately in reading and math using multiple-choice tests.[5] Test scores from the assessment were again put onto a common international scale with mean 500 and SD 100.

[4] This section draws heavily on the information provided in the official PASEC 2019 report (PASEC 2020).

[5] For example, a question evaluating the second skill dimension in math asked the teacher to assume that she gave a student the task to write down the figure 'five thousand three hundred and twenty six' in number format, and that the student's answer was 500030026. The teacher should then decide which of the following multiple-choice options best described the source of the student's error: (a) the student failed to read the numbers correctly, (b) the student does not know the number board well, (c) the student transformed each word separately to a number, and (d) there is no logic behind the student's answer. 50 percent of teachers across all countries picked the correct answer (c), whereas 30 percent picked answer (b).

3.2 Variable definitions

The dependent variables in our regressions are individual-level student test scores in reading and math. Test scores for each subject are reported as five plausible values, which are random draws from a posterior distribution. To obtain unbiased coefficient estimates, we use the averages of these five values as outcomes.[6] The two key explanatory variables are country-level averages of teacher subject knowledge and didactic skills scores in reading and math. For ease of interpretation, we transform student and teacher scores into z-scores by subtracting 500 and dividing by 100. In this way, we can interpret coefficients as the percent change in student achievement, measured in terms of international standard deviations, associated with an increase in teacher skills by one international standard deviation.

In some of our regressions, we control for a range of student, teacher, and school characteristics. These variables are derived from information collected via questionnaires, which PASEC fielded to students, teachers, and principals alongside the tests. We proxy for families' socioeconomic status using the number of books at home, availability of electricity at home, and parents' literacy (as reported by the student). We also observe a variety of school characteristics, including enrollment, whether the school is private or public, whether the school practices multigrade teaching, and an infrastructure index that summarizes information on the availability of resources such as running water, electricity, and toilets. Finally, we observe two subject-specific measures of textbook availability: an indicator for whether the student has her own textbook in class and an indicator for whether she can bring this textbook home.

We also construct country-level measures of population-wide literacy and numeracy from external data sources. Data on literacy come from the World Bank's World Development Indicators and reflect the share of a country's adult population that can both read and write (World Bank 2021*b*).[7] As internationally comparable data on numeracy are not available for most Sub-Saharan African countries, we use heaping patterns in self-reported age to construct a proxy. The intuition of this measure is as follows: in low-education settings, people might not be aware of their exact age, for example because they are unable to calculate the difference in years between the current year and their birth year. When asked about their age, they therefore tend to systematically round off to the nearest multiple of five or ten. This generates patterns of age heaping in population-wide survey data, which previous research has shown to be a good proxy for basic numeracy (see e.g. Duncan-Jones 2002, A'Hearn et al. 2009, Baten et al. 2014). We follow this research and create an index that captures age heaping patterns in the nationally representative Afrobarometer surveys, which are available for 11 of the 14 countries participating in PASEC 2019. We provide full details of this procedure in Appendix B. For our analysis, we standardize both literacy and numeracy to have mean zero and SD one across countries.

[6] Plausible values are used in most international student assessments, including PISA. For a detailed discussion of plausible values, see Jerrim et al. (2017).

[7] Data on literacy are recorded yearly but are not available for all years for every country. We always use the year closest to 2019 for which data are available (the earliest year we use is 2016).

3.3 Sample selection and descriptive statistics

Our sample consists of all sixth-grade students who participated in PASEC 2019. Table 1 reports summary statistics for this sample, which comprises 62,934 students in 3,298 schools in 14 countries. Students are 12.76 years old on average. Reflecting the low-income context, 21 percent of students do not have any literate parent, 35 percent do not have electricity at home, and 56 percent do not have any books at home. 26 percent of students attend a private school, and most of their schools are located in rural areas. Table 1 also reveals that as is usual in survey data, information on some control variables is missing for some students. In our regressions, we impute missing values on controls at the sample mean and include separate dummies for missing values on each control variable in order not to unnecessarily reduce sample size.[8]

To get a sense of the level and variation of student achievement and teacher quality, we present means of raw student and teacher scores separately by country in Table A1. Student achievement varies considerably between countries: average reading scores range from 451 in Chad to 645 in Gabon and average math scores range from 438 in Chad to 558 in Senegal. To put these figures into perspective, PASEC defined learning levels that compare a student's performance to the knowledge expected from sixth-grade students. According to this scale, students with scores above 517 (520) are considered to have 'sufficient' knowledge in reading (math). Notably, a large number of students does not reach this minimum level: average student scores are below this threshold in reading (math) in seven (nine) countries, which confirms previous findings of low average levels of learning in Sub-Saharan Africa (see World Bank 2018).

Teacher subject knowledge also differs substantially between countries: average reading scores range from 407 in Madagascar to 589 in Ivory Coast and average math scores range from 419 in Chad to 571 in Benin, differences which correspond to more than 1.5 international standard deviations. PASEC defined levels of proficiency that further facilitate the interpretation of these numbers. The scale considers scores below 393 (456) in reading (math) to require 'special attention and targeted training,' as teachers with such scores possess at most the very minimum knowledge for teaching. Notwithstanding the substantial cross-country variation in teacher scores, teachers in all countries score, on average, above this threshold in reading. In contrast, teachers in four countries do not reach scores above this cutoff in math. Moving beyond these average scores, teacher-level data reveal that across all participating countries, 16 percent (35 percent) of teachers score below the threshold in reading (math) (PASEC 2020). This finding corroborates previous results showing a lack of basic subject knowledge among teachers in Sub-Saharan Africa (see e.g. Bietenbeck et al. 2018, Bold et al. 2019).

Finally, Table A1 reveals cross-country differences in teacher didactic skills that are of similar magnitude to the gaps in subject knowledge: average scores range from 430 in the Republic of Congo to 579 in Ivory Coast in reading and from 409 in Guinea to 570 in Togo in math. Both

[8]Table 1 reveals that student achievement z-scores do not exactly have mean zero and SD one as might have been expected. The reason is that the international student achievement scale was normalized to the population of sixth-grade students assessed in PASEC 2014 and that additional countries participated in PASEC 2019 and student achievement changed over time. We confirmed that normalizing scores to have exactly mean zero and SD one in our sample yields estimates that are very similar to the ones presented in this paper. Similarly, Table 1 shows that the standard deviations of all teacher skills variables are less than one. This is because the scales of these variables were normalized within the sample of teachers, whereas we show summary statistics for the sample of students.

ranges correspond to roughly 1.5 international standard deviations. While no proficiency scale was developed for didactic skills, the PASEC 2019 report documents that teachers performed poorly on the test: across all countries, correct-answer rates for individual questions ranged from 43 percent to 55 percent in reading and from 23 percent to 55 percent in math (PASEC 2020). These results point towards the importance of distinguishing different dimensions of teacher quality: although teachers in all countries reach, on average, a sufficient level of subject knowledge in reading, they appear to have considerable difficulties to adapt this knowledge for teaching purposes as measured by the didactic skills test.

4 Empirical strategy

As a benchmark, we first estimate the following basic education production function separately for reading and math:

$$Y_{iksc} = \tilde{\alpha} + \tilde{\beta}_1 TSK_{kc} + \tilde{\beta}_2 TDS_{kc} + X_{isc}\tilde{\gamma}_1 + X_{iksc}\tilde{\gamma}_2 + X_{sc}\tilde{\gamma}_3 + P_{kc}\tilde{\gamma}_4 + \tilde{\varepsilon}_{iksc}. \tag{1}$$

Here, i denotes students, k denotes subjects, s denotes schools, and c denotes countries. Y_{iksc} is the subject-specific student test score, TSK_{kc} is the average teacher subject knowledge score in the country and subject, and TDS_{kc} is the average teacher didactic skills score. X_{isc} is a vector of subject-invariant student characteristics, such as gender, age and family background, X_{iksc} is a vector of subject-specific controls at the student level, such as textbook availability in class, and X_{sc} is a vector of subject-invariant controls at the school level. P_{kc} denotes subject-specific population skills at the country level. ε_{iksc} is the error term.

Despite the large number of control variables included in the specification in Equation 1, estimates of $\tilde{\beta}_1$ and $\tilde{\beta}_2$ are unlikely to reflect the causal effects of teacher subject knowledge and teacher didactic skills on student test scores due to omitted variable bias. For example, countries which place a greater value on education might have both higher skilled teachers and higher parental support for education. Since we cannot control for parental support, this would likely bias upward the estimated effects of teacher skills. Alternatively, countries in which home environments are less conducive to learning might employ higher skilled teachers in order to compensate for this disadvantage, biasing estimates downward. More generally, Section 2 shows that the countries and education systems in our sample differ on numerous dimensions, many of which could be correlated with both teacher skills and student achievement.

To overcome omitted variable bias, in our main regressions we exploit the fact that PASEC assessed both students and their teachers in two subjects. In particular, we ask whether *differences* in teacher skills between reading and math are systematically related to *differences* in student test scores between these subjects. This implies that we identify the effects of teacher skills only from within-student variation. We implement this method by pooling the data for reading and math and adding student fixed effects λ_i to the specification in Equation 1. This eliminates all subject-invariant controls from the regression:

$$Y_{iksc} = \beta_1 TSK_{kc} + \beta_2 TDS_{kc} + X_{iksc}\gamma_2 + P_{kc}\gamma_4 + \lambda_i + \phi_k + \varepsilon_{iksc}. \tag{2}$$

The student fixed effects in Equation 2 ensure that estimates of the effects of teacher skills are not biased by omitted variables whose influence does not differ between reading and math. Moreover, the fact that teacher skills are measured at the country level means that sorting of students to schools based on subject-specific factors is not an issue. The specification also accounts for two remaining potential sources of bias: first, countries with higher skilled teachers in math relative to reading might systematically emphasize the importance of numeracy over literacy. Such systematic emphasis could influence student achievement via channels other than teacher skills and would likely be reflected in unequal skills in the population. We therefore control for population-wide numeracy and literacy (P_{kc}). Second, countries with higher skilled teachers in a given subject might also have better physical school resources in that subject. We therefore control for the availability of textbooks (X_{iksc}), which have long been considered a key resource for learning in the context of Sub-Saharan Africa (e.g. Fredriksen & Brar 2015). Finally, we include a subject dummy (ϕ_k) in the regression in order to account for differences in average achievement between reading and math.

The specification in Equation 2 identifies the causal effects of teacher subject knowledge and teacher didactic skills under the assumption that conditional on student fixed effects and subject-specific controls, there are no other unobserved determinants of student achievement that correlate with teacher skills. The regression moreover assumes that the effect of teacher skills is equal in both subjects, an assumption that we provide evidence in favor of in Section 5. Finally, note that any cross-subject spillover effects of teacher skills on student achievement are netted out in this specification. Since such spillovers would likely be positive, this implies that our estimates reflect a lower bound of the true impact of teacher skills.

We estimate the specifications in Equations 1 and 2 using ordinary least squares. We weight all regressions using the student sampling weights provided with the PASEC 2019 data and give each country the same weight. We cluster standard errors by country and base our inference on wild cluster bootstrapped p values in order to account for the relatively low number of clusters (14 countries) in our sample (Cameron & Miller 2015, Abadie et al. 2017). To implement this method, we use Stata's –boottest– package (Roodman et al. 2019). We confirmed that this method of inference is conservative: when using conventional clustering instead, p values for the coefficients on teacher skills are always smaller.

5 Results

5.1 Benchmark estimates

Table 2 presents estimates based on the specification in Equation 1. There is a strong positive association between each dimension of teacher quality and student achievement in reading and math. Column 1 shows that in a regression without any controls, a one SD increase in subject knowledge is associated with a 0.59 SD (0.45 SD) rise in student reading (math) scores. Similarly, column 2 reveals that a one SD increase in teacher didactic skills is associated with a 0.60 SD (0.40 SD) rise in reading (math) scores. Column 3 shows that population-wide literacy and numeracy are also positively related to student achievement, although the corresponding coefficients are smaller and not statistically significant at conventional levels.

Columns 4 and 5 show results from regressions which include the full set of control variables. Compared to the uncontrolled regressions in columns 1 and 2, the coefficients on teacher skills are substantially reduced: a one SD increase in teacher subject knowledge is associated with a 0.39 SD (0.33 SD) rise in reading (math) scores, and a one SD increase in teacher didactic skills is associated with a 0.38 SD (0.26 SD) rise in reading (math) scores. A noteworthy pattern in these regressions is that the coefficients for reading and math are quite similar in magnitude. This finding supports the assumption of equal effects across subjects that we make in the student fixed-effects model in Equation 2.

Column 6 shows results from regressions in which both dimensions of teacher quality are included simultaneously. In these horse race specifications, only subject knowledge is strongly positively associated with student achievement, whereas the coefficient on didactic skills is much smaller for math and even negative for reading. Investigating this change in results, we find that subject knowledge and didactic skills are very highly correlated at the country level, with correlation coefficients of 0.95 for reading and 0.91 for math. This could lead to multicollinearity issues, and indeed we find substantially inflated standard errors in the regressions in column 6. The results in this final column should therefore be interpreted with caution.

5.2 Main student fixed-effects estimates

As discussed in Section 4, the benchmark estimates in Table 2 are unlikely to reflect the causal effects of teacher subject knowledge and teacher didactic skills due to omitted variables, which could bias the coefficients upward or downward. Therefore, we now turn to the student fixed-effects estimates based on Equation 2, which account for the influence of any subject-invariant confounders. Table 3 presents the results. Column 1 shows that in a regression without any further controls, a one SD increase in teacher subject knowledge is estimated to raise student achievement by 0.71 SD. Similarly, column 2 shows that a one SD increase in teacher didactic skills is estimated to raise student test scores by 0.58 SD.

Column 3 shows a small positive effect of population skills on student achievement, which is however not statistically significant at conventional levels. Columns 4 and 5 add population skills and textbook availability as controls to the regressions from columns 1 and 2. This does not change the coefficients on teacher subject knowledge and teacher didactic skills much, which suggests that unobserved subject-specific factors do not bias our results (Altonji et al. 2005).

Column 6 includes both dimensions of teacher quality in the same regression. Unlike in the benchmark regressions, we can disentangle the effects of subject knowledge and didactic skills in this specification because the *between-subject differences* of these variables are not as highly correlated at the country level (correlation coefficient of 0.59). The results show that the estimated effect of a one SD rise in teacher subject knowledge is almost identical to that found in columns 1 and 4 at 0.69 SD. In contrast, the estimated effect of a one SD rise in teacher didactic skills is substantially lower compared to columns 2 and 5 at 0.07 SD and not statistically significant at conventional levels. The results from this horse race regression thus reveal that differences in teacher subject knowledge, but not differences in teacher didactic skills, explain cross-country gaps in student learning in francophone Sub-Saharan Africa.

To gain an understanding of how much differences in teacher subject knowledge between countries matter, consider the case of Chad, which has among the lowest student achievement and teacher subject knowledge in reading and math (see Table A1). Our results suggest that if Chad's teachers had the same reading knowledge as teachers in Ivory Coast (the country with the highest average teacher subject knowledge in reading), its students would score 579 points on the reading test on average, earning them third rank among the fourteen countries participating in PASEC 2019. Similarly, if Chad's teachers had the same math knowledge as teachers in Togo (the country with the highest average teacher subject knowledge in math), its students would score 534 points on the math test on average, earning them fifth rank.[9] Thus, differences in teacher subject knowledge explain a large share of the observed differences in student achievement between Chad and the highest-achieving countries.

To get a more general picture of how important differences in teacher quality are for international learning gaps, we ran a country fixed-effects regression corresponding to our main specification in Equation 2. In particular, we first collapsed our data into country-subject cells. We then regressed student achievement on the two teacher skills variables and country fixed effects. The estimated effects of teacher skills in this regression were (mechanically) identical to those in column 6 of Table 3, and the between-country R-squared was 0.36. Thus, teacher subject knowledge and teacher didactic skills together account for more than a third of the international variation in average student achievement.

We next compare the estimates in Table 3 to our benchmark estimates and to findings in the previous literature. Compared to the benchmark results in column 4 of Table 2, the estimated effect of teacher subject knowledge in the student fixed-effects model is substantially larger. This suggests that the benchmark estimates are negatively confounded by unobserved student, school, or country characteristics. Our estimate seems large also when compared to results from the previous literature: for example, Hanushek et al. (2019) find that a one SD increase in teacher cognitive skills raises student achievement by 0.11 SD in a sample of mostly high-income countries, and Bietenbeck et al. (2018) and Bold et al. (2019) find that a one SD increase in teacher subject knowledge raises student achievement across several Sub-Saharan African countries by 0.03 SD and 0.07 SD, respectively.

The fact that our estimate is larger than those found in previous research is partly a statistical artifact: since our aim is to explain cross-country differences, we measure teacher skills and student achievement in terms of international standard deviations. In contrast, the above-mentioned papers normalize teacher skills at the student level. An implication is that a one-SD change in teacher skills in those papers corresponds to less than a one-SD change in terms of international standard deviations, leading to smaller point estimates.[10] Another potential reason for our comparatively large estimate is that we measure teacher subject knowledge that closely overlaps with the knowledge students are assessed on, whereas Hanushek et al. (2019) measure more general teacher cognitive skills, for example. Ultimately, we are unable to pin down the

[9] To calculate these figures, we multiply the difference in actual teacher subject knowledge in reading between Ivory Coast and Chad (589.3 – 420.8) and the difference in actual subject knowledge in math between Togo and Chad (556.1 – 419.3) by the estimated effect of teacher subject knowledge in column 6 of Table 3 (0.704) and add the result to the actual achievement of students in Chad in the corresponding subject.

[10] Consistent with this argument, Table 1 shows standard deviations of around 0.5 for the teacher skills variables in our student-level data, and consequently the estimated effect of teacher subject knowledge drops to about half the size if we normalize teacher skills to have SD one within our sample.

exact reason why our estimate is comparatively large, and this difference might be mostly due to differences in context.

5.3 Linearity, complementarity, and heterogeneity

We now summarize the results of several additional analyses based on the student-fixed effects specification. First, Figure 1 visualizes the regression in column 6 of Table 3 using binned scatter plots and reveals that the effect of teacher subject knowledge is roughly linear. Neither this effect nor the null effect of teacher didactic skills appears to be driven by outliers. Second, we tested for potential complementarity between teacher subject knowledge and teacher didactic skills by including an interaction term in the regression. The results showed no evidence of such complementarity: the point estimate for the interaction term was 0.01, although the confidence interval was wide and included economically meaningful effects. We also explored whether teacher skills and access to textbooks are complementary but found no evidence to this effect.

Third, we examine whether the effect of teacher skills varies by characteristics of students, schools, and countries. Table 4 presents estimates for various subsamples. Columns 1 and 2 show that the effect of teacher subject knowledge is slightly larger for girls than for boys. Columns 3 and 4 reveal that it is also larger for students who report having at least some books at home, a proxy for relatively higher wealth. In contrast, didactic skills appear to be more important for students without any books at home, although the estimated effect of 0.09 SD is not statistically significant at conventional levels. Columns 5 and 6 show that the impact of teacher subject knowledge is larger for students attending public schools. Finally, columns 7 and 8 reveal that subject knowledge matters more in countries with higher GDP.

5.4 Robustness

We now present the results from two robustness checks, which substantiate the validity of our findings. First, we verify that our estimates are not driven by any single country. Specifically, we re-run the headline regression of column 6 in Table 3 while excluding countries from the sample one by one. The results are presented in Table A2 and show that the estimated coefficients on teacher subject knowledge and teacher didactic skills in these restricted samples are very similar to the main estimates. Second, we ensure that our results are not sensitive to the imputation of missing values in our measure of population numeracy. Specifically, we draw on the Multiple Indicator Cluster Survey (MICS) to supplement our data set with self-reported age data for the three countries not participating in the Afrobarometer surveys. We then apply the procedure described in Appendix B to calculate the numeracy index for all 14 countries in our sample. Table A3 reports results from regressions in which we use this alternative measure of numeracy as control and reveals that our estimates are robust to this change.

6 Conclusion

School enrollment in Sub-Saharan Africa has risen dramatically in the past two decades. However, children in this region are learning very little in school, which limits the positive effect this educational expansion has on growth. In this paper, we focus on the large cross-country differences behind this low average level of learning and examine to what extent differences in teacher quality can explain international learning gaps in the region.

Our analysis builds on novel data from PASEC 2019, which let us observe student achievement and two dimensions of teacher quality, subject knowledge and didactic skills, for 14 francophone Sub-Saharan African countries. To identify the causal effect of teacher skills, we exploit variation between reading and math in a student fixed-effects model. Our main finding is that teacher subject knowledge has a large positive effect on student achievement, whereas the effect of didactic skills is comparatively small and not statistically significant at conventional levels. Together, the two dimensions of teacher quality explain 36 percent of the cross-country variation in average student achievement in our sample.

Our results show that teacher quality, and especially teacher subject knowledge, is a crucial driver of cross-country differences in learning. This is an important insight for policymakers in Sub-Saharan Africa who are trying to boost learning in schools, as it shows that there is a large payoff to recruiting more knowledgeable teachers. But given widespread difficulty to fill open positions, such improved recruitment might not be feasible in the short term – in any case, it would take many years for a change in recruitment practices to have an appreciable effect on average student learning. This renders in-service training for current teachers a potentially attractive alternative policy for boosting subject knowledge. While the quality of most teacher training programs is poor, recent research offers insights into how such programs can be designed to effectively boost teacher skills and student learning (see World Bank 2018).

References

Abadie, A., Athey, S., Imbens, G. W. & Wooldridge, J. (2017), When Should You Adjust Standard Errors for Clustering?, NBER Working Paper No. 24003.

A'Hearn, B., Baten, J. & Crayen, D. (2009), 'Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital', *The Journal of Economic History* **69**(3), 783–808.

Altonji, J. G., Elder, T. E. & Taber, C. R. (2005), 'Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools', *Journal of Political Economy* **113**(1), 151–184.

Baten, J., Crayen, D. & Voth, H.-J. (2014), 'Numeracy and the Impact of High Food Prices in Industrializing Britain, 1780–1850', *Review of Economics and Statistics* **96**(3), 418–430.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010), 'Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress', *American Educational Research Journal* **47**(1), 133–180.

Bergbauer, A. B., Hanushek, E. A. & Woessmann, L. (2021), 'Testing', *Journal of Human Resources* (forthcoming).

Bietenbeck, J. & Collins, M. (2020), 'New Evidence on the Importance of Instruction Time for Student Achievement on International Assessments', *Journal of Applied Econometrics* (forthcoming).

Bietenbeck, J., Piopiunik, M. & Wiederhold, S. (2018), 'Africa's Skill Tragedy: Does Teachers' Lack of Knowledge Lead to Low Student Performance?', *Journal of Human Resources* **53**(3), 553–578.

Bold, T., Filmer, D., Molina, E. & Svensson, J. (2019), The Lost Human Capital: Teacher Knowledge and Student Achievement in Africa, World Bank Policy Research Working Paper No. 8849.

Cameron, A. C. & Miller, D. L. (2015), 'A Practitioner's Guide to Cluster-Robust Inference', *Journal of Human Resources* **50**(2), 317–372.

Cueto, S., León, J., Sorto, M. A. & Miranda, A. (2017), 'Teachers' Pedagogical Content Knowledge and Mathematics Achievement of Students in Peru', *Educational Studies in Mathematics* **94**(3), 329–345.

Duncan-Jones, R. (2002), *Structure and Scale in the Roman Economy*, Cambridge University Press.

Fredriksen, B. & Brar, S. (2015), *Getting Textbooks to Every Child in Sub-Saharan Africa: Strategies for Addressing the High Cost and Low Availability Problem*, World Bank Publications.

Hanushek, E. A. (2013), 'Economic Growth in Developing Countries: The Role of Human Capital', *Economics of Education Review* **37**, 204–212.

Hanushek, E. A., Kinne, L., Lergetporer, P. & Woessmann, L. (2021), 'Patience, Risk-Taking, and Human Capital Investment across Countries', *Economic Journal* (forthcoming).

Hanushek, E. A., Link, S. & Woessmann, L. (2013), 'Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA', *Journal of Development Economics* **104**, 212–232.

Hanushek, E. A., Piopiunik, M. & Wiederhold, S. (2019), 'The Value of Smarter Teachers: International Evidence on Teacher Cognitive Skills and Student Performance', *Journal of Human Resources* **54**(4), 857–899.

Hanushek, E. A. & Rivkin, S. G. (2006), 'Teacher Quality', *Handbook of the Economics of Education* **2**, 1051–1078.

Hanushek, E. A. & Rivkin, S. G. (2012), 'The Distribution of Teacher Quality and Implications for Policy', *Annual Review of Economics* **4**(1), 131–157.

Hanushek, E. A. & Woessmann, L. (2011), 'The Economics of International Differences in Educational Achievement', *Handbook of the Economics of Education* **3**, 89–200.

Hanushek, E. A. & Woessmann, L. (2012*a*), 'Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation', *Journal of Economic Growth* **17**(4), 267–321.

Hanushek, E. A. & Woessmann, L. (2012*b*), 'Schooling, Educational Achievement, and the Latin American Growth Puzzle', *Journal of Development Economics* **99**(2), 497–512.

Hendricks, L. & Schoellman, T. (2018), 'Human Capital and Development Accounting: New Evidence from Wage Gains at Migration', *The Quarterly Journal of Economics* **133**(2), 665–700.

Hill, H. C., Rowan, B. & Ball, D. L. (2005), 'Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement', *American Educational Research Journal* **42**(2), 371–406.

Jerrim, J., Lopez-Agudo, L. A., Marcenaro-Gutierrez, O. D. & Shure, N. (2017), 'What Happens when Econometrics and Psychometrics Collide? An Example Using the PISA Data', *Economics of Education Review* **61**, 51–58.

Jones, B. F. (2014), 'The Human Capital Stock: a Generalized Approach', *American Economic Review* **104**(11), 3752–77.

Lavy, V. (2015), 'Do differences in schools' instruction time explain international achievement gaps? evidence from developed and developing countries', *The Economic Journal* **125**(588), F397–F424.

Marshall, J. H., Chinna, U., Nessay, P., Hok, U. N., Savoeun, V., Tinon, S. & Veasna, M. (2009), 'Student Achievement and Education Policy in a Period of Rapid Expansion: Assessment Data Evidence from Cambodia', *International Review of Education* **55**(4), 393–413.

Metzler, J. & Woessmann, L. (2012), 'The Impact of Teacher Subject Knowledge on Student Achievement: Evidence from Within-Teacher Within-Student Variation', *Journal of Development Economics* **99**(2), 486–496.

Ngo, F. J. (2013), 'The Distribution of Pedagogical Content Knowledge in Cambodia: Gaps and Thresholds in Math Achievement', *Educational Research for Policy and Practice* **12**(2), 81–100.

PASEC (2020), 'PASEC 2019 Qualité des Systèmes Éducatifs en Afrique Subsaharienne Francophone: Performances et Environnement de l'Enseignement-Apprentissage au Primare'.

Rockoff, J. E., Jacob, B. A., Kane, T. J. & Staiger, D. O. (2011), 'Can You Recognize an Effective Teacher When You Recruit One?', *Education Finance and Policy* **6**(1), 43–74.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G. & Webb, M. D. (2019), 'Fast and Wild: Bootstrap Inference in Stata Using Boottest', *The Stata Journal* **19**(1), 4–60.

Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N. & Miller, J. L. (2013), 'The Influence of Teachers' Knowledge on Student Learning in Middle School Physical Science Classrooms', *American Educational Research Journal* **50**(5), 1020–1049.

Shulman, L. (1987), 'Knowledge and Teaching: Foundations of the New Reform', *Harvard Educational Review* **57**(1), 1–23.

Shulman, L. S. (1986), 'Those Who Understand: Knowledge Growth in Teaching', *Educational Researcher* **15**(2), 4–14.

Woessmann, L. (2016), 'The Importance of School Systems: Evidence from International Differences in Student Achievement', *Journal of Economic Perspectives* **30**(3), 3–32.

World Bank (2018), *World Development Report 2018: Learning to Realize Education's Promise*, The World Bank.

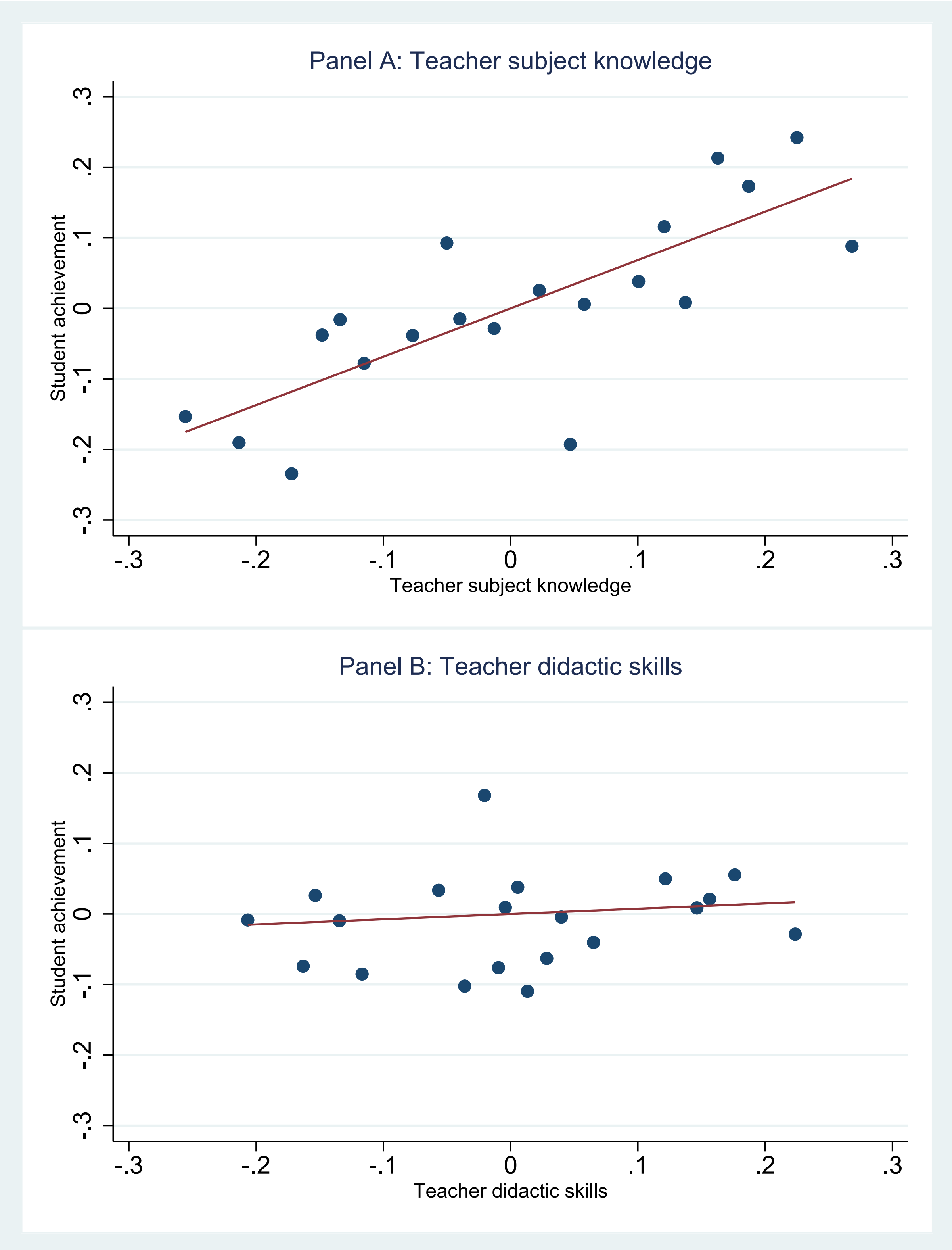
World Bank (2021*a*), GDP per Capita (current USD). https://data.worldbank.org/indicator/NY.GDP.PCAP.CD. Data accessed on May 25, 2022.

World Bank (2021*b*), Literacy Rate, Adult Total (Percent of People Ages 15 and Above). https://data.worldbank.org/indicator/SE.ADT.LITR.ZS. Data accessed on May 25, 2022.

World Bank (2021*c*), Poverty Headcount Ratio at USD 1.90 a day (2011 PPP) (Percent of Population). https://data.worldbank.org/indicator/SI.POV.DDAY?locations=ZG. Data accessed on May 25, 2022.

World Bank (2021*d*), School Enrollment, Primary (Percent Gross) in Low Income Countries. https://data.worldbank.org/indicator/SE.PRM.ENRR. Data accessed on May 25, 2022.

Figure 1
Effects of teacher skills on student achievement



Notes: The figure shows estimates of the effects of teacher subject knowledge (Panel A) and teacher didactic skills (Panel B) on student achievement in reading and math. Regressions are based on the student fixed-effects specification in column 6 of Table 3. To construct these plots, we first residualize student achievement and teacher skills on the controls included in this specification. We then group the residualized teacher skills variables into 20 equal-sized bins and plot the mean residualized achievement for each bin. The regression line in each plot is based on the underlying student-level data and thus visualizes the regression in column 6 of Table 3.

Table 1
Summary statistics

	Mean	SD	Number of students
Student achievement			
Reading z-score	0.20	1.08	62,934
Math z-score	–0.02	0.90	62,934
Teacher skills			
Reading subject knowledge z-score	0.00	0.59	62,934
Math subject knowledge z-score	0.00	0.51	62,934
Reading didactic skills z-score	0.00	0.52	62,934
Math didactic skills z-score	0.00	0.53	62,934
Student characteristics			
Male	0.51	0.50	62,917
Age	12.76	1.73	62,738
Electricity available at home	0.65	0.48	59,222
Student feels hungry in school	0.40	0.49	58,180
Books at home:			
No books	0.56	0.50	57,450
Enough to fill one shelf	0.32	0.47	57,450
Enough to fill two shelves	0.07	0.26	57,450
Enough to fill a bookcase	0.04	0.19	57,450
Literacy Parents:			
Illiterate	0.21	0.41	57,216
One parent literate	0.35	0.48	57,216
Both parents literate	0.43	0.50	57,216
School characteristics			
Private school	0.26	0.44	59,692
Infrastructure index	50.00	10.00	61,131
Enrollment	47.81	40.89	62,934
Multigrade school	0.25	0.43	61,112
School location:			
Town	0.36	0.48	60,767
Suburbs of town	0.09	0.29	60,767
Big village	0.30	0.46	60,767
Small village	0.25	0.43	60,767
Textbook availability			
Own reading textbook in class	0.73	0.45	59,873
Own math textbook in class	0.62	0.48	59,395
Can bring reading textbook home	0.73	0.45	42,495
Can bring math textbook home	0.74	0.44	35,965
Population skills			
Literacy	0.00	1.00	62,934
Numeracy	0.00	1.00	49,805

Notes: The table shows means and standard deviations and the number of students observed with each variable for the 62,934 students included in the analysis sample.

Table 2
Benchmark estimates

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Reading			
Teacher subject knowledge	0.584**			0.394**		0.770
	[0.016]			[0.031]		[0.102]
Teacher didactic skills		0.600**			0.380**	−0.460
		[0.029]			[0.033]	[0.149]
Population skills (literacy)			0.098	−0.036	−0.021	−0.044
			[0.528]	[0.725]	[0.821]	[0.609]
No. of observations	62,934	62,934	62,934	62,934	62,934	62,934
R-squared	0.104	0.083	0.008	0.369	0.359	0.373
			Panel B: Math			
Teacher subject knowledge	0.454**			0.327		0.242
	[0.020]			[0.129]		[0.691]
Teacher didactic skills		0.396**			0.255	0.085
		[0.010]			[0.163]	[0.845]
Population skills (numeracy)			0.091	0.004	0.016	0.006
			[0.479]	[0.960]	[0.884]	[0.929]
No. of observations	62,934	62,934	62,934	62,934	62,934	62,934
R-squared	0.066	0.055	0.008	0.260	0.258	0.260
Controls included:						
Student characteristics	no	no	no	yes	yes	yes
School characteristics	no	no	no	yes	yes	yes
Textbook availability	no	no	no	yes	yes	yes

Notes: The table shows estimates of regressions of student achievement in reading (Panel A) and math (Panel B) on teacher subject knowledge and teacher didactic skills. Regressions are based on the specification in Equation 1. All regressions use student sampling weights and give equal weight to all countries. For a detailed list of student characteristics, school characteristics, and textbook availability controls included in some of the regressions, see Table 1. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron & Miller (2015) and account for clustering at the country level. * p<0.10, ** p<0.05, *** p<0.01.

Table 3
Student fixed-effects estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Teacher subject knowledge	0.710**			0.720**		0.685***
	[0.040]			[0.017]		[0.008]
Teacher didactic skills		0.579*			0.576**	0.074
		[0.056]			[0.033]	[0.628]
Population skills			0.042	0.063	0.048	0.062
			[0.525]	[0.343]	[0.502]	[0.351]
No. of observations	125,868	125,868	125,868	125,868	125,868	125,868
R-squared	0.271	0.177	0.111	0.278	0.182	0.279
Controls included:						
Textbook availability	no	no	no	yes	yes	yes

Notes: The table shows estimates of the effects of teacher subject knowledge and teacher didactic skills on student achievement in reading and math. Regressions are based on the student fixed-effects specification in Equation 2. All regressions use student sampling weights and give equal weight to all countries. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron & Miller (2015) and account for clustering at the country level. * p<0.10, ** p<0.05, *** p<0.01.

Table 4
Heterogeneous effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gender		Books at home		School type		Country's GDP	
	Girls	Boys	No books	Some books	Public	Private	Below median	Above median
Teacher subject knowledge	0.723***	0.649***	0.633***	0.731**	0.670***	0.537*	0.475	1.011**
	[0.006]	[0.002]	[0.002]	[0.040]	[0.005]	[0.089]	[0.203]	[0.016]
Teacher didactic skills	0.071	0.083	0.093	−0.024	0.065	0.077	−0.012	−1.031
	[0.680]	[0.510]	[0.496]	[0.914]	[0.668]	[0.788]	[0.969]	[0.172]
No. of observations	61,814	64,020	69,214	45,686	93,358	26,026	74,130	51,738
R-squared	0.315	0.248	0.226	0.412	0.233	0.421	0.132	0.444

Notes: The table shows estimates of the effects of teacher subject knowledge and teacher didactic skills on student achievement in reading and math, separately for different groups of students. In columns 1-6, the sample is split by students' gender, the number of books students report having at home, and school type as indicated in the column headers. The number of observations in these specifications is lower than that in column 6 of Table 3 due to missing information on these variables. Columns 7 and 8 split the sample by countries' GDP per capita, with GDP data downloaded in March 2022 from the World Bank website: https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD. Regressions are based on the student fixed-effects specification in Equation 2. All regressions use student sampling weights and give equal weight to all countries. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron & Miller (2015) and account for clustering at the country level. * p<0.10, ** p<0.05, *** p<0.01.

A Additional tables

Table A1

Means of student achievement and teacher skills by country

	BDI	BEN	BFA	CIV	CMR	COD	COG	GAB	GIN	MDG	NER	SEN	TCD	TGO
Student achievement														
Reading	489.95	585.74	551.48	502.80	529.71	472.69	542.01	644.67	502.93	459.49	471.02	575.90	450.88	496.09
Math	546.01	533.82	547.17	453.97	488.13	462.09	489.11	554.61	482.27	468.32	461.80	557.58	437.78	495.39
Teacher skills														
Reading subject knowledge	461.50	548.40	550.40	589.30	542.70	420.90	467.30	548.50	449.70	407.30	484.50	561.80	420.80	546.80
Math subject knowledge	536.30	571.10	532.20	548.30	517.50	431.00	430.70	501.20	437.00	485.30	484.00	550.30	419.30	556.10
Reading didactic skills	457.00	536.20	543.10	578.90	539.40	437.40	430.10	540.70	460.40	450.50	487.40	572.50	436.90	529.60
Math didactic skills	493.90	551.70	558.30	533.40	518.80	411.10	442.80	521.40	409.00	479.90	518.30	553.30	438.10	570.10

Notes: The table shows means of student achievement and teacher skills by country. Country abbreviations: BEN = Benin, BFA = Burkina Faso, BDI = Burundi, CMR = Cameroon, CIV = Ivory Coast, GAB = Gabon, GIN = Guinea, MDG = Madagascar, NER = Niger, SEN = Senegal, TGO = Togo, COD = Democratic Republic of Congo, COG = Republic of Congo, TCD = Chad.

Table A2

Robustness: leaving one country out at a time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Country excluded:													
	BDI	BEN	BFA	CIV	CMR	COD	COG	GAB	GIN	MDG	NER	SEN	TCD	TGO
Teacher subject knowledge	0.544	0.726***	0.764***	0.693***	0.690***	0.632***	0.650**	0.545***	0.649***	0.858*	0.739***	0.684**	0.695***	0.712***
	[0.107]	[0.000]	[0.002]	[0.007]	[0.005]	[0.005]	[0.013]	[0.004]	[0.009]	[0.082]	[0.002]	[0.010]	[0.004]	[0.007]
Teacher didactic skills	0.067	0.107	−0.024	0.105	0.079	0.149	0.114	0.093	0.182	0.039	−0.024	0.099	0.074	0.005
	[0.710]	[0.350]	[0.888]	[0.550]	[0.597]	[0.497]	[0.598]	[0.418]	[0.311]	[0.735]	[0.886]	[0.569]	[0.635]	[0.985]
No. of observations	116,052	118,222	112,870	118,246	116,422	117,108	118,018	120,008	120,218	116,352	114,710	118,204	116,220	113,634
R-squared	0.262	0.299	0.300	0.266	0.272	0.293	0.258	0.213	0.296	0.300	0.298	0.289	0.292	0.293

Notes: The table shows estimates of the effects of teacher subject knowledge and teacher didactic skills on student achievement in reading and math. In each specification, one of the fourteen countries is excluded from the sample as indicated in the column header. For country abbreviations, see the notes to Table A1. Regressions are based on the student fixed-effects specification in Equation 2. All regressions use student sampling weights and give equal weight to all countries. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron & Miller (2015) and account for clustering at the country level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A3
Robustness: alternative measure of population numeracy

	(1)	(2)	(3)	(4)
Teacher subject knowledge	0.733**			0.703***
	[0.024]			[0.002]
Teacher didactic skills		0.576**		0.064
		[0.037]		[0.684]
Population skills	0.092	0.037	0.028	0.091
	[0.522]	[0.668]	[0.820]	[0.505]
No. of observations	125,868	125,868	125,868	125,868
R-squared	0.279	0.180	0.112	0.280

Notes: The table shows estimates of the effects of teacher subject knowledge and teacher didactic skills on student achievement in reading and math. Population skills are measured as described in Section 3.2 and use data from the Multiple Indicator Cluster Surveys (MICS) to replace missing values of countries not surveyed in the Afrobarometer. Regressions are based on the student fixed-effects specification in Equation 2, and include survey fixed-effects. All regressions use student sampling weights and give equal weight to all countries. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron & Miller (2015) and account for clustering at the country level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

B International data on numeracy

As internationally comparable data on numeracy skills are not available for most Sub-Saharan African countries, we construct a proxy based on heaping patterns in self-reported age. This measure exploits that people in low-education settings might not be aware of their exact age, for example because they are unable to calculate the number of years from the birth year to the current year. Therefore, when asked about their age, they tend to systematically round off to the nearest multiple of five or ten. This leads to patterns of age heaping in population-wide survey data, which have been shown to be a good proxy for basic numeracy (see e.g. Duncan-Jones 2002, A'Hearn et al. 2009, Baten et al. 2014).

We propose a variation of the Whipple Index to construct an internationally comparable measure of numeracy. The intuition of the index is to compare the frequency of reported ages ending in multiples of five to their predicted frequency in the absence of age heaping. Without any age heaping, we would expect 20% of the population having ages that end in multiples of five, assuming a uniform distribution of terminal digits. Hence, we calculate the ratio between the number of self-reported ages ending in five or ten and the expected number with uniformly distributed terminal digits:

$$W = \frac{\sum(n_{15} + n_{30} + ... + n_{60})}{\frac{1}{5}\sum_{i=15}^{64} n_i} \times 100 \tag{3}$$

For the assumption of uniformly distributed ages to hold, we need to restrict the age range to an interval in which each terminal digit occurs an equal number of times. Therefore, we only consider a subset of the population with reported ages between 15-64. The resulting index ranges between 0 and 500, with 500 indicating perfect heaping, which means all reported ages end in multiples of five, and 0 indicating perfect anti-heaping, which means that no individual reported an age ending in five or ten. A value of 100 thus indicates that exactly 20% of the population report their age to end in five or ten, implying that there exists no age heaping. We transform the index to range between 0 and 1, so that its value can be interpreted as the share of people that correctly report their age:

$$\widetilde{W} = \left(1 - \frac{W - 100}{400}\right) \times 100 \tag{4}$$

For our main analysis, we compute the index using self-reported age data from nationally representative Afrobarometer surveys. In particular, we use data from the seventh wave of the Afrobarometer which was conducted in 2019, the same year as the PASEC assessment we use in our analysis. For Burundi, we use data from the 2016 wave as it did not participate in more recent waves. Data for the Democratic Republic of Congo, Chad and Congo is missing because these three countries have not been surveyed in any wave of the Afrobarometer. In our main analysis, we impute the missing values for these three countries at the sample mean of our numeracy index across all other countries. In a robustness check, we draw on additional data from the Multiple Indicator Cluster Survey (MICS) to compute the index for the three countries missing in the Afrobarometer.