Economics of Education Review 96 (2023) 102437

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Teacher quality and cross-country differences in learning in francophone Sub-Saharan Africa☆

Jan Bietenbeck [a,b,c,d,*], Natalie Irmert [a], Mohammad H. Sepahvand [a,e]

[a] *Lund University, Sweden*
[b] *CESifo, Germany*
[c] *DIW Berlin, Germany*
[d] *IZA, Germany*
[e] *European Center for Advanced Research in Economics and Statistics (ECARES), Université libre de Bruxelles, Belgium*

ARTICLE INFO

JEL classification:
I21
I25
O15

Keywords:
International learning gaps
Teacher quality
Sub-Saharan Africa

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 averages of teacher subject knowledge and teacher 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 teacher didactic skills is comparatively small but imprecisely estimated. Differences in teacher subject knowledge account for 37 percent of the variation in average student achievement across countries.

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. Hendricks & Schoellman, 2018; Jones, 2014; Schoellman, 2012). Especially cognitive skills, as measured by student performance on international standardized tests, are a crucial driver of economic growth (Hanushek, 2013; Hanushek & Woessmann, 2012a, 2012b). In Sub-Saharan Africa, 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 educational expansion has on growth (World Bank, 2018). Importantly, this low general level of learning masks substantial 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 (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 high- and low-income countries (e.g. Bietenbeck et al., 2018; Metzler & Woessmann, 2012; Rockoff et al., 2011). Correlational evidence also links didactic skills, which describe teachers' ability to adapt subject knowledge for teaching purposes, to student achievement (e.g. Baumert et al., 2010; Hill et al., 2005; 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.

☆ A previous version of this paper was circulated under the title "Teacher Subject Knowledge, Didactic Skills, and Student Learning in Francophone Sub-Saharan Africa." We thank Luca Repetto, Simon Wiederhold, and the audiences at the spring 2022 Copenhagen Education Network Workshop, ESPE 2022, AEDE 2022, and EALE 2022 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.

* Correspondence to: Lund University, Department of Economics, P.O. Box 7080, 220 07 Lund, Sweden.
E-mail address: jan.bietenbeck@nek.lu.se (J. Bietenbeck).

https://doi.org/10.1016/j.econedurev.2023.102437
Received 19 September 2022; Received in revised form 14 June 2023; Accepted 29 June 2023
Available online 1 August 2023

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. We merge student-level data on achievement to country-level averages of teacher subject knowledge and didactic skills in each subject (sub-national data on teacher quality are not available to us). Our main regressions 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 very imprecisely estimated. These estimates imply that differences in teacher subject knowledge account for 37 percent of the variation in average student achievement across countries.

In additional analyses, we test whether the effects of teacher quality differ by student and country characteristics. We find that the effect of teacher subject knowledge is similar for girls and boys, but that the effect is larger in richer countries. We also explore potential nonlinearities and complementarities in the impacts of subject knowledge and didactic skills. While we find no evidence to this effect, our results are limited by the fact that we only observe country-level averages of the two teacher variables.

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 (Bietenbeck & Collins, 2023; Lavy, 2015), student testing (Bergbauer et al., 2021), and time and risk-taking preferences (Hanushek et al., 2022) 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 determinants of international learning gaps in low-income countries.

Our results also add to the literature on teacher quality in low-income countries. One strand of this literature shows that teacher quality, as captured by teacher value added, is of great importance for student learning (e.g. Araujo et al., 2016; Azam & Kingdon, 2015; Bau & Das, 2020). Interestingly, most commonly observed teacher characteristics, such as education, are not related to value added. In contrast, subject knowledge, which is one of the dimensions of teacher quality that we study, correlates with value added. Another strand of the literature studies interventions aimed at boosting teacher quality, such as in-service teacher training (Ashraf et al., 2021) and performance pay programs (Mbiti et al., 2019; Muralidharan & Sundararaman, 2011). These studies find substantial learning gains for students from such interventions. A major difference between these studies and our paper is our focus on explaining differences in learning between rather than within countries. Moreover, we provide some of the first evidence on the effect of didactic skills, which is a little-studied potential dimension of teacher quality.[2]

2. Education in francophone Sub-Saharan Africa

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, 2021c).[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, 2021c). Similarly, GDP per capita ranges from 239 USD in Burundi to 6,882 USD in Gabon (World Bank, 2021a). Countries also experience very different levels of violence and social unrest, with 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, 2021d). 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.

[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. Baumert et al., 2010; Cueto et al., 2017; Hill et al., 2005; Marshall et al., 2009; Ngo, 2013; Sadler et al., 2013). 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.

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 (*Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN*) 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.

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

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.

3.2. Cross-country differences in student achievement and teacher quality

To get a sense of the level and variation of student achievement and teacher quality, we present averages of student and teacher scores separately by country in Table 1. 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 1 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 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

[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).

Table 1
Average student achievement, teacher subject knowledge, and teacher didactic skills by country and subject.

	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
Difference	−56.07	51.92	4.31	48.83	41.58	10.60	52.90	90.05	20.66	−8.84	9.23	18.32	13.10	0.70
Teacher subject knowledge														
Reading	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	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
Difference	−74.80	−22.70	18.20	41.00	25.20	−10.10	36.60	47.30	12.70	−78.00	0.50	11.50	1.50	−9.30
Teacher didactic skills														
Reading	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	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
Difference	−36.90	−15.50	−15.20	45.50	20.60	26.30	−12.70	19.30	51.40	−29.40	−30.90	19.20	−1.20	−40.50

Notes: The table shows average 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.

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.[6]

3.3. Sample construction and summary statistics

We base our analysis on the PASEC 2019 data made available to researchers. These are student-level data, which contain information on achievement in reading and math and socio-demographic characteristics, but no information on teacher quality. We merge these data to averages of teacher subject knowledge and teacher didactic skills at the country-by-subject level, which we extract from the official PASEC 2019 report (PASEC, 2020). Observing teacher quality at the country level is sufficient to estimate its effect on cross-country differences in average student achievement. However, it limits our possibility to detect non-linear effects and complementarities, a caveat that we discuss in more detail in Section 5.4.

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.[7] 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, our regression coefficients capture the impact of a one SD increase in teacher skills on student achievement, measured in terms of international standard deviations.

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, 2021b).[8] 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. A'Hearn et al., 2009; Baten et al., 2014; Duncan-Jones, 2002). 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.

Our sample consists of all sixth-grade students who participated in PASEC 2019. Table 2 reports summary statistics for this sample, which comprises 62,934 students 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 2 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

[6] Interestingly, teacher subject knowledge and teacher didactic skills in both reading and math are highly correlated with measures of state capacity, which describes the ability of states to implement and enforce policies. For example, the correlation coefficient of teacher subject knowledge in reading and the state capacity index (Hanson & Sigman, 2021) is 0.75 (it is 0.79 for the government effectiveness index (World Bank, 2019), 0.28 for the quality of government index (Dahlberg et al., 2023) and 0.52 for the state fragility index (Center for Systematic Peace, 2023). As the existing measures are imperfect in capturing all aspects of state capacity, measures of teacher quality could be considered as a supplement to existing indices in future research.

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

[8] 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).

Table 2
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.

mean and include separate dummies for missing values on each control variable in order not to unnecessarily reduce sample size.[9]

4. Empirical strategy

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

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

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. P_{kc} denotes subject-specific population skills at the country level. X_{iksc} contains the two indicators of subject-specific textbook availability. X_{isc} is a vector of subject-invariant student characteristics, such as gender, age and family background, and X_{sc} is a vector of subject-invariant controls at the school level. ε_{iksc} is the error term.

The two dimensions of teacher quality that we observe, TSK_{kc} and TDS_{kc}, enter the regression in Eq. (1) via the function $f(\cdot)$. This reflects the fact that ex ante, the exact nature of their impact is unclear. For example, teacher subject knowledge might matter for student learning only up to a certain threshold level of knowledge. Another possibility is that subject knowledge and didactic skills are complements in educational production. In our initial regressions, we nevertheless model the two variables as linear and additively separable. This choice is motivated by the fact that (i) previous studies on teacher subject knowledge find no evidence of non-linear effects (Bietenbeck et al., 2018; Hanushek et al., 2019), and (ii) the extant literature on teacher effects tends to model factors such as teacher experience, education, and test scores as additively separable (e.g. Aslam & Kingdon, 2011; Clotfelter et al., 2010; Rockoff et al., 2011). The linear and additively separable specification is also parsimonious, which is particularly important given that we exploit international variation from only 14 countries. Nonetheless, we relax this restriction and allow for more flexible functional forms of $f(\cdot)$ later on.

With $f(TSK_{kc}, TDS_{kc}; \tilde{\beta}) = \tilde{\beta}_1 TSK_{kc} + \tilde{\beta}_2 TDS_{kc}$, our benchmark specification reads:

$$Y_{iksc} = \tilde{\alpha} + \tilde{\beta}_1 TSK_{kc} + \tilde{\beta}_2 TDS_{kc} + \tilde{\gamma}_1 P_{kc} + X_{iksc}\tilde{\gamma}_2 + X_{isc}\tilde{\gamma}_3 + X_{sc}\tilde{\gamma}_4 + \tilde{\varepsilon}_{iksc}. \tag{2}$$

Despite the large number of control variables included in this regression, 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 revealed 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.[10] We implement this method by pooling the data for reading and math and adding student fixed effects λ_i to the specification in Eq. (2). We also add a subject dummy ϕ_k in order to account for differences in average achievement between reading and math, and we drop all subject-invariant controls from the regression:

$$Y_{iksc} = \beta_1 TSK_{kc} + \beta_2 TDS_{kc} + \gamma_1 P_{kc} + X_{iksc}\gamma_2 + \lambda_i + \phi_k + \varepsilon_{iksc}. \tag{3}$$

The student fixed effects in Eq. (3) ensure that the estimated effects of teacher subject knowledge and teacher didactic skills are not biased by omitted variables whose influence does not differ between reading

[9] Table 2 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 2 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.

[10] Within-student between-subject variation has been widely used to estimate teacher effects in the literature; see, for example, Dee (2007), Metzler and Woessmann (2012), Bietenbeck (2014), Bietenbeck et al. (2018), Hanushek et al. (2019), and Bold et al. (2019).

and math, such as students' general academic ability. 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 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).

One potential concern with the specification in Eq. (3) is that estimates could still be biased due to unobserved subject-specific ability that correlates with teacher quality. Importantly, because teacher quality is measured at the country level, such subject-specific ability would have to manifest itself at the country level too. But country-level subject-specific ability among students intuitively correlates with population-wide skills, which we control for in the regression. We acknowledge, however, that this control may be imperfect, and that in the end our specification cannot account for all potential biases due to subject-specific factors. In Section 5 below, we compare the coefficients β_1 and β_2 from regressions with and without subject-specific controls P_{kc} and X_{iksc} in order to get an insight into the importance of such unobserved subject-specific confounders (Altonji et al., 2005).

Another potential concern with the student-fixed effects model is the assumption that the effects of teacher subject knowledge and teacher didactic skills are the same in reading and math. The previous literature has provided evidence in favor of this assumption: for example, Bietenbeck et al. (2018) and Hanushek et al. (2019) show that in regressions of student achievement on teacher skills, the coefficients are not statistically different between specifications focusing on reading and math. Similarly, Araujo et al. (2016) show that an encompassing measure of teacher quality has the same effect on student learning in math and language in Ecuador, and Bau and Das (2020) show that teacher value added is comparable for math and language scores in Pakistan. In Section 5 below, we provide evidence suggesting that the equal effects assumption holds also in our setting by showing that the coefficients on teacher skills in the benchmark regressions are similar for reading and math.[11]

We estimate the specifications in Eqs. (2) and (3) 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 14 country clusters in our sample (Abadie et al., 2017; Cameron & Miller, 2015). 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 3 presents estimates based on the specification in Eq. (2). 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 0.58 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. Columns 3 and 4 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.

Column 5 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 5. The results in this final column should therefore be interpreted with caution.

The last two rows in Table 3 show p values from tests of the null hypothesis of equal coefficients on teacher subject knowledge and teacher didactic skills in both subjects. The null cannot be rejected in any of the columns, indicating that the coefficients in the reading and math specifications are not statistically different from each other. This finding is in line with the results from the previous literature and supports the assumption of the student-fixed effects model that the impacts of teacher quality are the same in these two subjects.[12]

5.2. Main student fixed-effects estimates

As discussed in Section 4, the benchmark estimates in Table 3 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 Eq. (3), which account for the influence of any subject-invariant confounders. Table 4 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. Columns 3 and 4 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 5 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 3 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 4 at 0.07 SD and very imprecisely estimated. The main take-away from Table 4 is thus that differences in teacher subject knowledge explain cross-country gaps in 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,

[11] Note that if there are cross-subject spillover effects of teacher skills on student achievement, these are netted out in the student-fixed effects specification. Since such spillovers would likely be positive, this implies that our estimates reflect a lower bound of the true impact of teacher skills.

[12] The finding that the association of teacher skills and student achievement does not differ between subjects is in line with the previous literature in economics an are among.

Table 3
Benchmark estimates.

	(1)	(2)	(3)	(4)	(5)
			Panel A: Reading		
Teacher subject knowledge	0.584**		0.394**		0.770*
	[0.016]		[0.031]		[0.096]
Teacher didactic skills		0.600**		0.380**	−0.460
		[0.029]		[0.041]	[0.149]
No. of observations	62,934	62,934	62,934	62,934	62,934
R-squared	0.104	0.083	0.369	0.359	0.373
			Panel B: Math		
Teacher subject knowledge	0.454**		0.327		0.242
	[0.016]		[0.149]		[0.726]
Teacher didactic skills		0.396**		0.255	0.085
		[0.015]		[0.169]	[0.883]
No. of observations	62,934	62,934	62,934	62,934	62,934
R-squared	0.066	0.055	0.260	0.258	0.260
Controls included:					
Student characteristics	no	no	yes	yes	yes
School characteristics	no	no	yes	yes	yes
Textbook availability	no	no	yes	yes	yes
Population skills	no	no	yes	yes	yes
Tests of equal coefficients:					
p($TSK_{read} = TSK_{math}$)	0.339		0.718		0.373
p($TDS_{read} = TDS_{math}$)		0.293		0.521	0.357

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 Eq. (2). All regressions use student sampling weights and give equal weight to all countries. For a detailed list of the controls included in some of the regressions, see Table 2. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron and Miller (2015) and account for clustering at the country level. The last two rows show p values from tests of the null hypothesis of equal coefficients on teacher subject knowledge (TSK) and teacher didactic skills (TDS) in panels A and B, respectively. We test this hypothesis by stacking the observations for reading and math and running a regression in which the subject dummy is interacted with all right-hand-side variables. The p values reported in the table are from a test of the null that the interaction between the subject dummy and the respective teacher skill variable is zero. * p<0.10, ** p<0.05, *** p<0.01.

Table 4
Main student fixed-effects estimates.

	(1)	(2)	(3)	(4)	(5)
Teacher subject knowledge	0.710**		0.720**		0.685***
	[0.040]		[0.025]		[0.005]
Teacher didactic skills		0.579*		0.576**	0.074
		[0.056]		[0.035]	[0.628]
No. of observations	125,868	125,868	125,868	125,868	125,868
R-squared	0.271	0.177	0.278	0.182	0.279
Controls included	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 Eq. (3). All regressions use student sampling weights and give equal weight to all countries. Regressions in columns 3 to 5 include subject-specific population skills and textbook availability as controls. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron and Miller (2015) and account for clustering at the country level. * p<0.10, ** p<0.05, *** p<0.01.

which has among the lowest student achievement and teacher subject knowledge in reading and math (see Table 1). 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.[13] 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 estimate country-level regressions corresponding to the specification in column 5 of Table 4. Fig. 1 shows the results. The upper panel illustrates the effect of teacher subject knowledge on student achievement. To construct this plot, we first residualize teacher subject knowledge and student achievement on the control variables and teacher didactic skills. We then collapse our data into country-by-subject cells and plot between-subject differences in conditional student achievement against between-subject differences in conditional teacher subject knowledge. The effect of teacher subject knowledge, which is illustrated by the

[13] 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 4 (0.704) and add the result to the actual achievement of students in Chad in the corresponding subject.

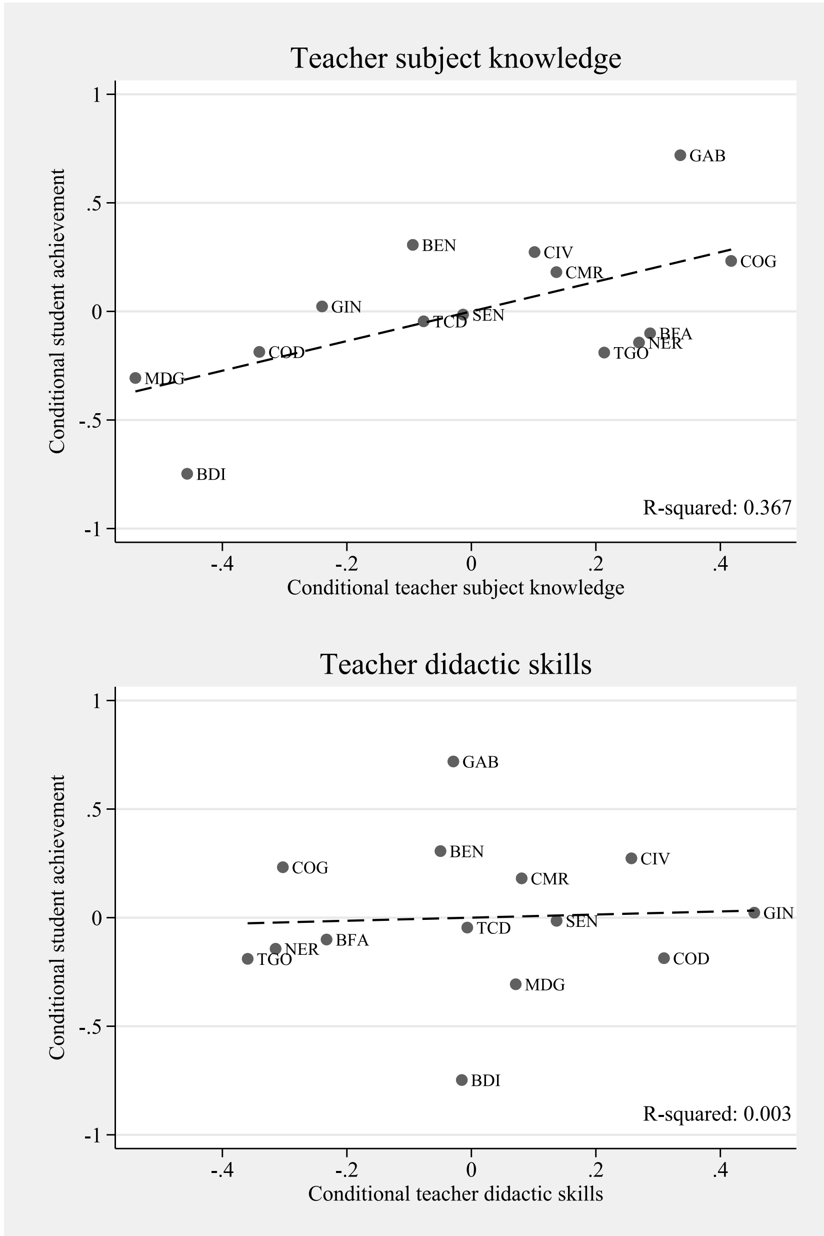

Fig. 1. Effects of teacher skills on student achievement.
Notes: The figure shows results from country-level regressions that are equivalent to the student-level regression in column 5 of Table 4. The top panel shows the effect of teacher subject knowledge, conditional on teacher didactic skills and control variables. The bottom panel shows the effect of teacher didactic skills, conditional on teacher subject knowledge and the controls. To construct these plots, we first residualize student achievement and the teacher skill variable on the controls and the respective other teacher skill. We then collapse the data at the country-by-subject level and plot between-subject differences in conditional student achievement against between-subject differences in conditional teacher skills. The regression line and R-squared value reported in each plot comes from the country-level regression of these differences.

dashed line in the plot, is mechanically identical to the one estimated in column 5 of Table 4. But the country-level regression additionally reveals that differences in teacher subject knowledge account for 37 percent of the international variation in average student achievement, as evidenced by the R-squared of 0.367. The lower panel of Fig. 1 shows the corresponding results for teacher didactic skills. As could be expected from the estimates in Table 4, differences in teacher didactic skills explain only a very small fraction of the international variation in average student achievement once differences in teacher subject knowledge are accounted for.

5.3. Discussion of effect size

We next compare the estimates in Table 4 to our benchmark estimates and to findings in the previous literature. Compared to the benchmark results in column 3 of Table 3, 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 of the effect of teacher subject knowledge 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. We confirm this intuition in Table A1, in which we re-standardize our teacher variables to have mean zero and SD one across students in our sample. With this standardization, which corresponds to the standardization used in the previous literature, the estimated effect of teacher subject knowledge drops to 0.36. One potential reason why this estimate is still larger than those in the previous literature is that we measure 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, however, we are unable to pin down the exact reason why our estimate is so large, and this might be mostly due to differences in context.

Our estimated positive effect of teacher didactic skills is in line with results from studies in educational science, which also link didactic skills to student achievement (e.g. Baumert et al., 2010; Hill et al., 2005). The associations found in these studies are, however, unlikely to capture causal effects since the underlying analyses do not account for the likely sorting of students between and within schools, which limits comparisons with our estimates. Conducting a randomized controlled trial in Uganda, Ashraf et al. (2021) find that an in-service pedagogy training increased student achievement by 0.5–0.8 SD, which is substantially larger than our estimate. One potential explanation why their effect is larger is that their intervention implies a fundamental change in the pedagogical concept, whereas our analysis exploits marginal differences in didactic skills. Notwithstanding this difference in context, we point out that our estimate of teacher didactic skills is very imprecisely estimated, which means that we cannot reject economically meaningful effects.

5.4. Exploring different functional forms

In our main estimates, we assume that the effects of teacher subject knowledge and teacher didactic skills are linear and additively separable. Table 5 shows estimates from specifications in which we relax this assumption. In columns 1–3, we add quadratic terms for both teacher skills variables. There is little evidence of such non-linearities: indeed, the coefficients on the linear terms in column 3 are almost identical to the ones in column 5 of Table 4, although the effects are now less precisely estimated. In column 4, we instead allow for complementarities between subject knowledge and didactic skills by including an interaction term. Again, there is no evidence of such teacher skill complementarities.

One caveat with these results is that they are based only on variation in average teacher quality across 14 countries. These averages likely mask substantial within-country differences in teacher quality, and as a consequence the variation we rely on does not cover the extremes of the teacher quality distribution. Therefore, if there are non-linearities or complementarities at very low or very high levels of teacher subject knowledge and teacher didactic skills, the regression in Table 5 cannot detect these.[14] With this caveat in mind, given that we do not find evidence of non-linearities and complementarities, we continue to use the linear and additively separable specification in the rest of the paper.

Table 5
Different functional forms of teacher quality.

	(1)	(2)	(3)	(4)
Teacher subject knowledge	0.732*		0.673	0.691**
	[0.073]		[0.236]	[0.017]
Teacher didactic skills		0.583**	0.087	0.067
		[0.025]	[0.690]	[0.592]
(Teacher subject knowledge)^2	0.074		−0.001	
	[0.814]		[0.999]	
(Teacher didactic skills)^2		0.182	0.097	
		[0.463]	[0.589]	
Linear interaction				0.047
				[0.878]
No. of observations	125,868	125,868	125,868	125,868
R-squared	0.279	0.189	0.281	0.279

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 Eq. (3) but allow for different functional forms of the main explanatory variables. All regressions include subject-specific population skills and textbook availability as controls and use student sampling weights, giving equal weight to all countries. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron and Miller (2015) and account for clustering at the country level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

5.5. Effect heterogeneity

We now examine whether the effect of teacher skills varies by characteristics of students, schools, and countries. Table 6 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.[15]

5.6. 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 rerun the headline regression of column 5 in Table 4 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 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.

[14] This being said, Bietenbeck et al. (2018) used teacher-level data on subject knowledge and did not find any evidence of non-linear effects.

[15] In unreported regressions, we also explored whether the effect of teacher skills differs by textbook availability but found no evidence to this effect. This is in contrast to the results by Mbiti et al. (2019), who find strong complementarities between providing teacher incentives and school resources in the context of Tanzania.

Table 6
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]	[0.008]	[0.001]	[0.039]	[0.005]	[0.058]	[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]	[0.524]	[0.485]	[0.905]	[0.680]	[0.751]	[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 5 of Table 4 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 Eq. (3). All regressions include subject-specific population skills and textbook availability as controls and use student sampling weights, giving equal weight to all countries. p values in brackets are based on the wild cluster bootstrap procedure suggested by Cameron and Miller (2015) and account for clustering at the country level. * p<0.10, ** p<0.05, *** p<0.01.

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 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: differences in teacher subject knowledge account for 37 percent of the cross-country variation in average student achievement in our sample. In contrast, the estimated effect of teacher didactic skills is positive but comparatively small once subject knowledge is accounted for, but very imprecisely estimated.

An important question is whether these estimates reflect causal effects. By including student fixed effects and using only between-subject variation for identification, we believe that our estimates account for the most important potential biases. However, our results do rely on two untestable assumptions: the absence of subject-specific confounders and equal effects across subjects. While we provide evidence in favor of these assumptions, we cannot entirely rule out that violations of them bias our estimates.

Taken together, our results suggest 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).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.econedurev.2023.102437.

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.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *Quarterly Journal of Economics*, *131*(3), 1415–1453.

Ashraf, N., Banerjee, A., & Nourani, V. (2021). *Learning to teach by learning to learn: Working paper*.

Aslam, M., & Kingdon, G. (2011). What can teachers do to raise pupil achievement? *Economics of Education Review*, *30*(3), 559–574.

Azam, M., & Kingdon, G. G. (2015). Assessing teacher quality in India. *Journal of Development Economics*, *117*, 74–83.

Baten, J., Crayen, D., & Voth, H.-J. (2014). Numeracy and the impact of high food prices in industrializing Britain, 1780–1850. *The Review of Economics and Statistics*, *96*(3), 418–430.

Bau, N., & Das, J. (2020). Teacher value added in a low-income country. *American Economic Journal: Economic Policy*, *12*(1), 62–96.

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* (in press).

Bietenbeck, J. (2014). Teaching practices and cognitive skills. *Labour Economics*, *30*, 143–153.

Bietenbeck, J., & Collins, M. (2023). New evidence on the importance of instruction time for student achievement on international assessments. *Journal of Applied Econometrics*, *38*(3), 423–431.

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.

Center for Systematic Peace (2023). State fragility index and matrix, time-series data, 1995-2018. https://www.systemicpeace.org/inscrdata.html. (Data accessed 24 April 2023).

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *Journal of Human Resources*, *45*(3), 655–681.

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.

Dahlberg, S., Sundström, A., Holmberg, S., Rothstein, B., Pachon, N. A., Dalli, C. M., & Meijers, Y. (2023). The quality of government basic dataset, version jan23. https://data.worldbank.org/indicator/SE.ADT.LITR.ZS. (Data accessed 24 April 2023).

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, *42*(3), 528–554.

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.

Hanson, J. K., & Sigman, R. (2021). Leviathan's latent dimensions: Measuring state capacity for comparative political research. *The Journal of Politics*, *83*(4), 1495–1510.

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. (2022). Patience, risk-taking, and human capital investment across countries. *The Economic Journal 132*(646), 2290–2307.

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. In *Handbook of the economics of education, vol. 2* (pp. 1051–1078). Elsevier.

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. In *Handbook of the economics of education, vol. 3* (pp. 89–200). Elsevier.

Hanushek, E. A., & Woessmann, L. (2012a). 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. (2012b). 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. *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), 3752–3777.

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.

Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., & Rajani, R. (2019). Inputs, incentives, and complementarities in education: Experimental evidence from Tanzania. *Quarterly Journal of Economics*.

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.

Muralidharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of Political Economy*, *119*(1), 39–77.

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*. PASEC Dakar, Senegal.

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.

Schoellman, T. (2012). Education quality and development accounting. *Review of Economic Studies*, *79*(1), 388–417.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–23.

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 (2019). Government effectiveness: Estimate [GE.EST]. https://databank.worldbank.org/source/worldwide-governance-indicators#. (Data accessed 24 April 2023).

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

World Bank (2021b). *Literacy rate, adult total (Percent of people ages 15 and above)*. The World Bank, https://data.worldbank.org/indicator/SE.ADT.LITR.ZS. (Data accessed 25 May 2022).

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

World Bank (2021d). *School enrollment, primary (Percent gross) in low income countries*. The World Bank, https://data.worldbank.org/indicator/SE.PRM.ENRR. (Data accessed 25 May 2022).