The Value of Smarter Teachers: International Evidence on Teacher Cognitive Skills and Student Performance*

Eric A. Hanushek, Marc Piopiunik, Simon Wiederhold§

November 26, 2016

Abstract

International differences in teacher quality are commonly hypothesized to be a key determinant of the large international student performance gaps, but lack of consistent quality measures has precluded testing this. Using unique assessment data, we construct country-level measures of teacher cognitive skills. We find substantial differences in teacher cognitive skills across countries, and these are strongly related to student performance. Results are supported by fixed-effects estimation exploiting within-country between-subject variation in teacher skills. A series of robustness and placebo tests highlight the systematic influence of teacher skills as distinct from overall differences among countries in the level of cognitive skills. Moreover, observed country variations in teacher cognitive skills are significantly related to differences in women's access to high-skilled occupations outside teaching and to salary premiums for teachers.

Keywords: teacher cognitive skills, student performance, international comparison, PIAAC, PISA

JEL classification: I20, H40, J20

---

* We would like to thank William Thorn, Veronica Borg, Vanessa Denis, and Francois Keslair for access to and help with the international PIAAC data. We further thank Lisa Leschnig and Guido Schwerdt for making their international data on central exit examinations available to us. We are also indebted to David Deming, Bernhard Enzi, Oliver Falck, Josh Goodman, Brian Jacob, Susanna Loeb, Daniel Münich, Paul Peterson, Jesse Rothstein, Johannes Schmieder, Guido Schwerdt, Dick Startz, Marty West, Ludger Woessmann, and participants at the CESifo Area Conference on the Economics of Education, the annual conference of the Verein für Socialpolitik, the RWI Research Network Conference on the "Economics of Education," the Applied Micro Research Seminar at CERGE-EI, the annual conference of the European Economic Association, the Third International PIAAC Conference, and seminars in Konstanz and Munich for providing valuable comments. Piopiunik is indebted to the Program on Education Policy and Governance (PEPG) at Harvard University, in particular to Paul Peterson and Antonio Wendland, for their support and hospitality during his research visit. Wiederhold is thankful for the hospitality provided by the Center for International Development at Harvard University, especially to Ricardo Hausmann, Ljubica Nedelkoska, and Frank Neffke. Wiederhold also gratefully acknowledges financial support from the Fritz Thyssen Foundation.

§ Hanushek: Hoover Institution, Stanford University, CESifo, and NBER, hanushek@stanford.edu; Piopiunik: Ifo Institute at the University of Munich and CESifo, piopiunik@ifo.de; Wiederhold: Ifo Institute at the University of Munich, CESifo, and Catholic University Eichstätt-Ingolstadt, wiederhold@ifo.de.

1. Overview

Numerous international assessment tests have shown that the cognitive skills of students differ greatly across developed countries. These differences take on considerable significance because the cognitive skills of the population have been shown to be an important driver of a country's long-run economic growth (e.g., Hanushek and Woessmann (2015)). But less considered is how the overall skills of a nation feed back into the skills of teachers. This paper investigates whether differences in cognitive skills of teachers across developed countries can help explain international differences in student performance.

Public discussions have emphasized the importance of teacher skills for improving student achievement. For example, a widely-cited McKinsey report on international achievement concludes that "the quality of an educational system cannot exceed the quality of its teachers" and then goes on to assert that "the top-performing systems we studied recruit their teachers from the top third of each cohort graduate from their school system." (Barber and Mourshed (2007), p. 16) In a follow-on report, Auguste, Kihn, and Miller (2010) note that the school systems in Singapore, Finland, and Korea "recruit 100% of their teacher corps from the top third of the academic cohort," which stands in stark contrast to the U.S. where "23% of new teachers come from the top third." (p. 5) They then recommend a "top third+ strategy" for the U.S. educational system. We investigate the implications for student achievement of focusing policy attention on the cognitive skills of potential teachers.

Our analysis exploits unique data from the Programme for the International Assessment of Adult Competencies (PIAAC), which allow for the first time quantifying differences in teacher skills in numeracy and literacy across countries. These differences in teacher cognitive skills reflect, as we discuss below, both the overall level of cognitive skills of each country's population and where teachers are drawn from in each country's skill distribution.

Teacher cognitive skills differ widely internationally. For example, average numeracy and literacy skills of teachers in countries with the lowest measured skills (Chile and Turkey) are well below the skills of employed adults with just vocational education in Canada.[1] In contrast, the skills of teachers in countries with the highest measured skills (Japan and Finland) exceed the skills of adults with a master's or PhD degree in Canada.

The evidence from combining this information on teacher quality with student achievement suggests that differences in teacher cognitive skills are a significant contributor to international differences in student performance. Specifically, we use country-level measures of subject-specific teacher skills along with rich student-level micro data from the Programme for International Student

[1] We use Canada as a benchmark for the international skill comparison because the Canadian sample is by far the largest among all countries surveyed in PIAAC, allowing for a fine disaggregation of individuals by educational degree.

Assessment (PISA) to estimate the impact of teacher cognitive skills on student performance in math and reading across 31 developed economies.

We pursue three different strategies to investigate the impact of teacher cognitive skills. First, we estimate OLS models with extensive sets of control variables, including student and family background, general and subject-specific school inputs, institutional features of the school systems, and cross-country differences in educational inputs. Controlling for subject-specific parent cognitive skills, which can be approximated with the PIAAC data, allows us to account for the persistence of skills across generations and to distinguish between smart parents and smart teachers. Nevertheless, countries may still differ from one another in other, hard-to-observe ways.

To circumvent bias due to unobserved heterogeneity, our second approach exploits the performance of students and teachers across two different subjects, math and reading. This allows us to identify the effect of teacher cognitive skills using only variation between subjects, which directly controls for unobserved student-specific characteristics that similarly affect math and reading performance (e.g., innate ability or family background). At the same time, this within-student between-subject model also controls for all differences across countries that are not subject-specific, such as general education preferences or the nature of teacher labor markets.

The fixed-effects estimates may still be biased by subject-specific country differences; for example, some countries may particularly emphasize math skills while others may attach more importance to reading skills. However, if there are subject-specific educational preferences, numeracy or literacy skills in occupations other than teaching should also be higher. Our third strategy therefore controls for the cognitive-skill levels in other occupations (14 in total, e.g., managers, scientists and engineers, health professionals, business professionals) and, alternatively, for the skill level of all adults. This specification hence exploits within-country variation in cognitive skills, placing teachers in the country-specific distribution of numeracy or literacy skills instead of exploiting variation only across countries. We also test in a placebo-type analysis whether the cognitive-skill levels in other occupations are related to student performance when the relevant skills – those of teachers – are included.

Across empirical strategies, the results indicate a robust positive impact of teacher cognitive skills on student performance. In the OLS estimation with the full set of controls, we find that a one standard deviation (SD) increase in teacher cognitive skills is associated with about 0.15 SD higher student performance in math and 0.10 SD higher student performance in reading. The between-subject estimates are in the same range as the OLS coefficients. When exploiting within-country variation in cognitive skills, teacher cognitive skills always keep their strong impact on student performance even when controlling for the cognitive-skill levels in other occupations or the cognitive skills of all adults. Importantly, these other measures of cognitive skills are almost never significantly related to student performance when controlling for teacher skills.

Our results are robust to adding coarse measures of teachers' subject-specific pedagogical skills, suggesting that international differences in teacher cognitive skills do not merely reflect differences in pedagogical skills. Moreover, accounting for cross-country differences in the stage of economic development and in educational institutions such as central exit exams or controlling for continental fixed effects to address issues of divergent national cultures does not change the teacher-skill coefficients.

We also provide novel evidence about the determinants of differences in teacher cognitive skills across countries. Existing studies have shown a strong decline in teacher cognitive skills in the U.S. during the past decades. This decline has been explained with improving alternative opportunities for women in the labor market (Bacolod (2007)). Using the PIAAC data, we generalize the U.S. evidence to a broader set of countries, exploiting within-country changes across birth cohorts in the proportion of females working in high-skilled occupations. By observing multiple countries, we can more readily assess how female labor-market opportunities interact with teacher quality.

Higher shares of women working in high-skilled occupations other than teaching are significantly related to lower cognitive-skill levels of teachers. This suggests that international differences in women's opportunities to enter (other) high-skilled occupations provide part of the explanation for the observed variation in teacher cognitive skills across countries.

The PIAAC micro data also permit looking explicitly at whether teachers in each country tend to be paid above or below what would be expected (given their gender, work experience, and cognitive skills). We find considerable variation in the premiums paid to teachers, with Ireland paying teachers considerably above market and the United States and Sweden paying teachers considerably below market. These country-specific premiums are directly related to the cognitive skills of teachers across countries.

The paper proceeds as follows. Section 2 considers relevant prior research. Section 3 introduces the datasets and describes the computation of our measures of teacher and parent cognitive skills. Section 4 presents our identification strategies. Section 5 reports results on the impact of teacher cognitive skills on student performance in math and reading and provides robustness checks and placebo tests. Section 6 analyzes possible determinants of the cross-country differences in teacher cognitive skills, focusing on women's access to alternative high-skilled occupations and on teacher salaries. Section 7 considers policy trade-offs more directly, and Section 8 concludes.

2. Relevant Literature

Large numbers of studies investigate the determinants of student achievement within individual countries.[2] The clearest conclusion from this "educational production function" literature is that achievement reflects a combination of family background factors, school inputs, and institutional factors. However, these studies are better suited for within-country analysis and are not structured to explain differences in achievement across countries. In particular, all of these studies consider the impacts of school characteristics within a country's overall institutional structure – such as the amount of local decision-making authority at schools, the requirements for teacher certification, and the overall salary levels for teachers – and do not necessarily give an accurate picture of their impact under differing institutional structures.

A parallel literature on international differences in achievement builds on the comparative outcome data in existing international assessments (see Hanushek and Woessmann (2011)). One of the clearest explanatory factors from these international studies has been the importance of family background in explaining student achievement.[3] In contrast, specific conclusions about the impact of school resources have been much more limited. There has, for example, been considerable research on overall educational expenditures and on resource inputs such as class size, but the existing research has not identified these as being strong drivers of international differences in achievement.[4] The lack of findings on resources has led to a different set of international studies that focuses on the effects of institutional features of the school systems such as the degree of local decision making, the use of accountability systems, and direct rewards for personnel in the schools.[5]

The most convincing within-country studies show that teacher impacts on student reading and math performance differ greatly and that there is huge variation in teacher value-added (Hanushek and Rivkin (2012)).[6] But these results have not been very useful in addressing international achievement differences. First, the studies focus almost exclusively on the experience in the United

[2] See, for example, the reviews in Hanushek (2002) and Glewwe et al. (2013).

[3] For example, see the review in Björklund and Salvanes (2011) or the analysis in Woessmann et al. (2009).

[4] See Hanushek (2006) for a review of the effects of school resources and the international evidence in Hanushek and Woessmann (2011).

[5] For example, positive impacts have been estimated for school autonomy (especially in developed countries; cf. Hanushek, Link, and Woessmann (2013)) and for increased competition reflected in the share of privately operated schools (West and Woessmann (2010)). The range of institutional studies is assessed in Hanushek and Woessmann (2011).

[6] For a sample of the research into teacher effectiveness, see Rockoff (2004), Rivkin, Hanushek, and Kain (2005), Kane, Rockoff, and Staiger (2008), Chetty, Friedman, and Rockoff (2014), and the summaries in Hanushek and Rivkin (2010). As an indication of the magnitudes involved, Rivkin, Hanushek, and Kain (2005) estimate that the effect of a costly ten student reduction in class size is smaller than the benefit of moving up the teacher quality distribution by one standard deviation.

States. Second, they have not reliably described any underlying determinants of teacher value-added – and in particular any determinants that can be consistently measured across countries.

Moreover, within-country studies (going beyond just the value-added studies) have generally shown that the common measures of teacher differences – education, experience levels, and sources and nature of teacher preparation – are not consistently related to student achievement, raising questions about the reliance on these as indicators of teacher quality in international studies. In a closely related set of within-country and international studies, researchers have used measures of teacher salaries as proxies for teacher quality, implicitly assuming that higher-paid teachers have higher skills or are more motivated. Dolton and Marcenaro-Gutierrez (2011) construct a country panel with international student assessment tests in the period 1995–2006, showing that teacher salaries – both measured in absolute terms and relative to the average wages in a country – are positively associated with student performance even after controlling for country fixed effects. Related analysis has looked at the use of performance pay, and the international research has tended to find that pay incentives are effective in improving performance.[7] However, the within-country evidence again indicates that teacher salaries are a weak measure of teacher quality (see the overview by Hanushek and Rivkin (2006)).[8]

One general strand of research, largely focused on entry and exit from teaching, investigates the importance of alternative job opportunities for teacher quality.[9] Although these studies look just within the U.S., they suggest feasible approaches to international comparisons. Nagler, Piopiunik, and West (2015) exploit business cycle conditions at career start as a source of exogenous variation in the outside options of potential teachers, finding that teachers entering the profession during recessions are significantly more effective in raising student test scores than teachers who entered the profession during non-recessionary periods. Other work, which forms an important motivation for our study, focuses on the cognitive skills of teachers over time – a dimension of teacher quality that had received some support in prior estimation of educational production functions.[10] Bacolod (2007) documents a decrease in the academic quality (as measured by standardized test scores and

---

[7] For a review on teacher performance pay, see Leigh (2013). See also the international investigation of performance pay in Woessmann (2011).

[8] Challenging this general conclusion, Britton and Propper (2016) find positive effects of relative teacher pay on school productivity, exploiting regional variation in teachers' relative wages. Loeb and Page (2000) similarly relate regional variation in relative teacher wages to rates of educational attainment but also lack direct measures of teacher quality. We explore the country-level relationship between teacher wage premiums and teacher cognitive skills in Section 6.

[9] Early estimation of outside opportunities on teacher transitions is found in Dolton and van der Klaauw (1999), although the key issues were suggested long before in Kershaw and McKean (1962). An early investigation of how preparation for and entry into teaching are related to cognitive skills is found in Hanushek and Pace (1995).

[10] While not completely consistent, previous research has found cognitive skill of teachers (as measured by scores on achievement tests) to be perhaps the strongest proxy of an underlying dimension of teacher quality (see Eide, Goldhaber, and Brewer (2004); Hanushek and Rivkin (2006), and the summary in Hanushek (2003)).

undergraduate institution selectivity) of female teachers in the U.S. during the recent decades that coincided with the expansion of job opportunities for women. Corcoran, Evans, and Schwab (2004a, (2004b) show that the decline in measured teacher skills over the same period has been concentrated in the upper portion of the achievement distribution. Both suggest that women's opportunities to enter high-skilled occupations outside teaching are one determinant of the skill level of teachers in a country.

The analysis of varying skill levels of teachers in these studies has, however, not been linked directly to student performance. There have been some studies of measured teacher cognitive skills, particularly in the U.S., but these have relied generally on small and idiosyncratic data sets, and the results have been far from consistent (see Hanushek (2003)).[11]

3. International Comparative Data

A unique feature of this study is the application of new and consistent international data on cognitive skills of teachers to assess the role of cross-country differences in teacher cognitive skills in explaining student outcomes.

3.1 Teacher Cognitive Skills

Measured cognitive skills of teachers are derived from the Programme for the International Assessment of Adult Competencies (PIAAC) survey. Developed by the Organisation for Economic Co-operation and Development (OECD) and collected in 2011/2012 (Round 1) and in 2014/15 (Round 2), PIAAC tested various cognitive skill domains of more than 215,000 adults in 33 developed economies.[12] The target population of PIAAC was the non-institutionalized population aged 16-65 years, and samples included at least 5,000 participants in each country.

The survey provides rich information about demographic, educational, and occupational characteristics for each respondent. It was administered by trained interviewers either in the respondent's home or in a location agreed upon between the respondent and interviewer. The standard

---

[11] In two studies for developing countries, Metzler and Woessmann (2012) and Bietenbeck, Piopiunik, and Wiederhold (2015) show the relevance of teacher subject knowledge using individual-level teacher data. See also Harbison and Hanushek (1992) for the impact of measured teacher math skills on achievement in rural Brazil. Using a general, non-subject-specific, measure on cognitive abilities (based on a standard IQ test), Grönqvist and Vlachos (2016) find only a negligible impact of teacher cognitive skills on student achievement.

[12] We use 31 countries in our analysis: Australia, Austria, Belgium (Flanders), Canada, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, the Netherlands, Norway, Poland, the Russian Federation, the Slovak Republic, Spain, Sweden, the United Kingdom (England and Northern Ireland), and the United States (Round 1) as well as Chile, Greece, Indonesia, Israel, Lithuania, New Zealand, Singapore, Slovenia, and Turkey (Round 2). Cyprus, while participating in PIAAC, did not participate in PISA. In Indonesia, the PIAAC survey was only administered to the population in Jakarta. According to OECD (2013), data for the Russian Federation are preliminary, may still be subject to change, and are not representative of the entire Russian population because they do not include the population of the Moscow municipal area. Our results are not sensitive to dropping the Russian Federation from the sample.

survey mode was to answer questions on a computer, but respondents without computer experience could opt for a pencil-and-paper interview.[13]

After providing the background information, respondents took a battery of cognitive assessments. PIAAC assessments are designed to be valid cross-culturally and cross-nationally and to provide internationally comparable measures of adult skills. The assessments measure key cognitive and workplace skills needed to advance in the job and to participate in society in three domains: numeracy, literacy, and problem solving in technology-rich environments.[14] The test questions are often framed as real-world problems, such as maintaining a driver's logbook (numeracy domain) or selecting key information from a bibliographic search (literacy domain). PIAAC measures each of the skill domains on a 500-point scale.

We are particularly interested in the skills of teachers in each country. In the Public Use File, information on occupation is available only at the two-digit code in some countries (Germany, Ireland, Singapore, Sweden, and the United States), while a few other countries (Austria, Canada, Estonia, and Finland) do not publicly report any occupational code. For this study, however, we gained access through the OECD to the four-digit ISCO-08 (International Standard Classification of Occupations) codes for all countries, which allows us to identify teachers in fine categories.

We define teachers as all PIAAC respondents who report as current four-digit occupation code "primary school teacher", "secondary school teacher", or "other teacher" (which includes, for example, special education teachers and language teachers).[15] We exclude university professors and vocational school teachers since the vast majority of PISA students (15-year-olds) are still in secondary school and have therefore not been taught by these types of teachers. We also exclude pre-kindergarten teachers as the roles of this teacher group depend directly on the institutional structures of individual countries and may or may not be contributors to teaching students reading and math.[16]

---

[13] On average across countries, 70 percent of assessment participants took the computer-based assessment and 30 percent took the paper-based assessment. A field test suggests no impact of assessment mode (OECD 2013).

[14] PIAAC tests were conducted in the official language of the country of residence. In some countries, the assessment was also conducted in widely spoken minority or regional languages. Respondents could take as much time as needed to complete the assessment. *Literacy* is defined as the "ability to understand, evaluate, use and engage with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential," and *numeracy* is the "ability to access, use, interpret, and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life" (see OECD (2013) for more details). Because of our focus on students' reading and math performance, we do not use the PIAAC skills in the domain "problem solving in technology-rich environments." Moreover, five countries surveyed in PIAAC (Cyprus, France, Indonesia, Italy, and Spain) did not administer tests in this optional skill domain.

[15] Results are very similar if we drop the category "other teachers." We keep these teachers in the sample to increase sample size.

[16] For Australia and Finland we only have two-digit occupational codes and are unable to exclude pre-kindergarten teachers and university professors/vocational school teachers from our teacher sample. However, analysis of the countries where teachers are defined using the four-digit code indicates that teacher skills based on the four-digit code are very similar to those defined using the two-digit code: The correlation of both skill measures is 0.97 for numeracy and 0.95 for literacy. On average, numeracy (literacy) skills based on the two-digit code are only marginally higher (by 0.5 (0.1) PIAAC points) than the respective skills based on the four-digit codes. The average absolute value of these differences is

PIAAC does not allow us to identify the subject that a teacher is teaching, so we use the numeracy and literacy skills of all teachers tested in PIAAC. We focus on the country-level median of the teacher cognitive skills.[17] We weight individual-level observations with inverse sampling probabilities when computing country-specific teacher cognitive skills.

Table 1 reports summary statistics of the teacher cognitive skills in the 31 countries and in the pooled sample. The number of teachers in the national PIAAC samples ranges from 106 teachers in Chile to 834 teachers in Canada, with 207 teachers per country on average. (The sample size for Canada is substantially larger than for any other country surveyed in PIAAC because Canada oversampled in order to obtain regionally representative adult skills). Teachers in Finland and Japan perform best in both numeracy and literacy, while teachers in Chile and Turkey perform worst in both domains. The range of teacher numeracy skills across countries is 55 points, which amounts to one international individual-level standard deviation (55 points) in the full PIAAC sample; in literacy, the range of teacher skills of 60 points even exceeds one international standard deviation (50 points).[18] Teachers in the United States (284 points) perform worse than the average teacher in numeracy (292 points) but are slightly above the international mean in literacy (301 points vs. 295 points). Interestingly, the country ranking and the cross-country variation in teacher cognitive skills are similar to those of all prime-aged workers with full-time employment (see Table 1 in Hanushek et al. (2015)).[19] Also note that teacher numeracy skills are higher than teacher literacy skills in some countries, while the reverse is true in other countries. We will exploit this variation in domain-specific teacher skills in the fixed-effects model that uses only variation within countries and between subjects (see Section 5.3). Furthermore, both numeracy and literacy skills of teachers are completely unrelated to the number of teachers in the national PIAAC samples. For the econometric analysis, we standardize the teacher cognitive skills across the 31 countries to have mean zero and standard deviation one.

Figure 1 provides an illustration of the international variation in teacher cognitive skills. The figure arrays the median teacher numeracy and literacy skills across countries against the skills of adults in different educational groups within Canada, the country with the largest total sample. The literacy skills of the lowest-performing teachers (in Turkey and Chile) are well below the literacy skills of employed Canadian adults with only a vocational degree (278 points). Teachers in Italy,

---

only 2.1 points in numeracy and 1.9 points in literacy. Moreover, simultaneously excluding Australia and Finland from the analysis does not qualitatively change our results below.

[17] The country-level correlation between teacher median skills and mean skills is 0.98 for both numeracy and literacy. Moreover, all results are robust to using mean teacher skills instead of median teacher skills.

[18] The respective means in the full PIAAC sample are 259 points in numeracy and 268 points in literacy.

[19] Younger teachers have higher skills than older teachers in almost all countries in our sample. Also, male teachers have higher skills than female teachers, especially in numeracy. These patterns, however, are not specific to teachers, but are very similar among all college graduates in a country. Detailed results are available on request.

Russia, and Israel perform just as well as vocationally educated Canadians. Teachers in Canada, the Netherlands, and Sweden have skill levels similar to Canadian adults with a bachelor degree (306 points). The literacy skills of the best-performing teachers (in Japan and Finland) are even higher than the skills of Canadian adults with a master or doctoral degree (314 points). These comparisons, which look similar for numeracy skills, underscore the vast differences in teacher cognitive skills across developed countries.

Variations in teacher cognitive skills reflect both where teachers are drawn from the cognitive skill distribution of the population and where a country's overall cognitive-skill level falls in the world distribution. As most teachers have obtained a college degree (89 percent on average across all PIAAC countries), we expect teacher cognitive skills to fall above the median of the skill distribution of the entire adult population. Across all 31 countries, median teacher skills fall at the 68th (71st) percentile of the numeracy (literacy) skill distribution of all adults, ranging from the 53rd to the 81st percentile (see Table 1).

It is also illuminating to compare teacher cognitive skills with the skills of all college graduates in a country (see Figure 2). While median teacher cognitive skills fall near the middle of the 25th–75th percentile skill range of cognitive skills of college graduates in most countries, teachers come from the upper part of the skill distribution in some countries (e.g., Finland, Japan, and Singapore) and from the lower part of the college graduate skill distribution in other countries (e.g., Austria, Denmark, the Slovak Republic, and Poland).

From Table 1, teachers in Chile, France, Spain, and Turkey are drawn highest up from the country distributions of adult skills in numeracy and literacy. Despite having the highest measured cognitive skills, Finnish teachers are drawn from a lower part of the country's overall skill distribution, reflecting the fact that average cognitive skills in Finland are considerably larger than in France and Spain, and substantially larger than in Chile and Turkey. Or, harkening back to the argument that 100% of teachers in Korea and Singapore come from the top third of the academic cohort, the median Korean (Singaporean) teacher falls at the 72nd (72nd) percentile of the overall country distribution and at the 52nd (48th) percentile of the college graduate distribution in numeracy (see also Figure 2).[20]

Because the PIAAC tests are new and have not been fully validated, we have compared the PIAAC-based teacher cognitive skills with the numeracy and literacy skills of teachers in larger national datasets for the United States and Germany. These comparisons, described in Appendix A, support the overall validity of the estimates of teacher cognitive skills that are derived from PIAAC.

[20] This point about teacher skills was first made by Schleicher (2013).

3.2 Parent Cognitive Skills

Because the parents of the PISA students (henceforth "PISA parents") are not tested themselves in any skill domain, we use the PIAAC data to impute the numeracy and literacy skills of the PISA parents. We first construct a sample of adult PIAAC participants that could in principle be PISA parents. We then estimate the numeracy and literacy skills of the PISA parents from the PIAAC micro data on the basis of several common observable characteristics. Specifically, separately by country, we regress the numeracy/literacy skills of PIAAC adults aged 35–59 with children[21] on three characteristics: gender[22], education (3 categories), and number of books at home (6 categories).[23] These estimated coefficients with the same three characteristics of the PISA parents are then used to obtain predicted numeracy/literacy skills of each PISA parent at the individual family level. In the student-level analysis, we use the maximum skills of mother and father as a proxy for parent cognitive skills, although results are very similar if the average skill of mother and father is used instead.

Although the PIAAC-based parent skills are only coarse proxies for the true skills of PISA parents, controlling for the estimated cognitive-skill level of parents allows us to tackle several issues. First, since originally studied in the Coleman Report (Coleman et al. (1966)), it has been clear that the family and education in the home is important. Using parental cognitive skills adds a qualitative dimension to family influences over and above the common measures of the student's family background. More generally, student performance likely persists across generations, for example, because the quality of the education system or the valuation of education changes only slowly over time. Second, adding information about parent cognitive skills provides one way of separating teacher cognitive skills from the skills of the country's overall population.

Table A-1 presents summary statistics of parent skills in numeracy and literacy by country. Similar to teacher cognitive skills, parent cognitive skills differ greatly across countries, ranging (in numeracy) from 223 points in Chile to 308 points in Japan. Also, parent skills differ substantially within countries. On average, the difference between the minimum and maximum numeracy skill in a country is 115 points, or more than twice the international individual-level standard deviation.

---

[21] Individuals in this age bracket are potential parents of the 15-year-old PISA students since they were 17–44 years old when PISA students were born.

[22] We compute skills separately for PISA mothers and fathers because numeracy/literacy skills of women and men might differ. By predicting gender-specific skills, PISA students with single mothers, for example, are assigned only the skill level of women and not the average skill level of men and women.

[23] We collapsed the original 8 categories of the PIAAC education variable into 3 categories so that the education categories in PIAAC and PISA would exactly match. The 6 categories of the number of books at home variable are identical in PIAAC and PISA, so this variable was not modified. We use number of books at home in addition to educational degree, since this variable has been shown to be the single strongest predictor of student test scores (Woessmann (2003)). Sample sizes range from 1,074 adults in the Russian Federation to 11,933 adults in Canada with an average sample size of 2,693 adults per country (see Table A-1).

3.3 Student Performance and Further Control Variables

International data on student performance come from the Programme for International Student Assessment (PISA), conducted by the OECD.[24] PISA is a triennial survey that tests math and reading competencies of nationally representative samples of 15-year-old students, an age at which students in most countries are approaching the end of compulsory schooling. PISA contains both multiple-choice and open-answer questions and provides internationally comparable test scores. The tests emphasize understanding as well as flexible and context-specific application of knowledge, and hence they do not test curriculum-specific knowledge.

We use the two PISA cycles of 2009 and 2012 because the student cohorts in these two test cycles have largely been taught by the teacher cohorts tested between 2011 and 2015 in PIAAC. Student cohorts of earlier PISA cycles (2000, 2003, and 2006) have partially been taught by some PIAAC teachers, but teacher turnover would introduce additional error in the teacher skill measures for students in these earlier cycles. Another reason for combining PISA 2009 and 2012 is that students provide information about the instructional practices of their teachers only for the subject that is the focus in each round of PISA testing: reading in 2009 and math in 2012. From the survey information, we can compute country-specific indicators of instructional practice for reading (based on PISA 2009) and for math (based on PISA 2012). These instructional-practice indicators capture subject-specific pedagogical skills of teachers, which might be a potentially important confounding factor for teacher cognitive skills (see Section 5.2).

Table A-2 provides summary statistics of student performance and student characteristics. Student performance in math and reading differs widely across countries. Given that the learning progress in one school year is about 40 PISA points, the difference between the USA and Singapore is about two school years in math. The math performance gap is about three schools years between Singapore and Turkey and almost four years between Singapore and Chile, which is at the bottom of our sample of countries. International student performance differences in reading are somewhat less pronounced, but still substantial.

For the econometric analysis, we standardize student test scores at the student level across the 31 countries and the two PISA assessments to have mean zero and standard deviation one. As we are interested in differences across countries, each country receives the same total weight in each PISA cycle. Student characteristics (e.g., gender and migration status) and information about parents (e.g., education, occupation, and number of books at home) come from student background questionnaires.

---

[24] We rely on the PISA assessments instead of the alternative international test of Trends in International Mathematics and Science Study, or TIMSS (see Hanushek and Woessmann (2011)). PISA covers more PIAAC countries, and students participating in PISA were tested in both math and reading, while TIMSS only assessed math performance. Note, however, that math scores from TIMSS are strongly correlated with math scores from PISA at the country level.

In addition to parent cognitive skills, we use number of books at home, parents' highest educational degree, and parental occupation to control for family background (see Table A-3).

Based on student information, we can construct measures of weekly instructional time for both language and math classes.[25] Furthermore, school principals provide information on whether the school is public or private, city size, total number of students in the school, the lack of qualified math teachers and language teachers, and different types of autonomy (see Table A-4).

Country characteristics include variables that are direct educational measures, namely, cumulative educational expenditure per student between age 6 and 15 and school starting age. We also check the robustness of our results to including further country controls, for instance, GDP per capita to capture international differences in the state of development (see Table A-5).

4. Estimation Strategy

In the baseline OLS model, we estimate an international education production function of the following form:

$$A_{iksc} = \alpha + \lambda T_{kc} + \boldsymbol{F}_{isc}\boldsymbol{\beta}_1 + \boldsymbol{S}_{sc}\boldsymbol{\beta}_2 + \boldsymbol{C}_c\boldsymbol{\beta}_3 + \gamma_1 P_{iksc} + \boldsymbol{I}_{ksc}\boldsymbol{\gamma}_2 + \epsilon_{iksc}, \quad \textbf{(1)}$$

where A_{iksc} is the test score of student *i* in subject *k* (math or reading) in school *s* in country *c*. T_{kc} represents the median teacher cognitive skills in subject *k* in country *c*. $\boldsymbol{F}_{isc}$ is a vector of student-level variables measuring student and family background, $\boldsymbol{S}_{sc}$ is a vector of school-level characteristics, and $\boldsymbol{C}_c$ is a vector of country-level control variables. P_{iksc} contains student-level estimates of parents' numeracy and literacy skills, respectively, and $\boldsymbol{I}_{ksc}$ contains school-level variables measuring the shortage of qualified teachers and weekly instructional time in math and language classes. (See Tables A1–A5 for descriptive statistics for all control variables). ε_{iksc} is an error term, assumed to be mean zero. Throughout, we cluster standard errors at the country level because teacher skills do not vary within countries.[26]

Despite the rich set of control variables, interpreting the OLS estimate of λ as the causal effect of teacher cognitive skills on student performance is nonetheless problematic, because of the

---

[25] Following Lavy (2015), we aggregate this information across students to the school level.

[26] Recent research has shown that clustered standard errors can be biased downward in samples with a small number of clusters (e.g., Donald and Lang (2007); Cameron, Gelbach, and Miller (2008); Angrist and Pischke (2009); Imbens and Kolesar (2012)). Although there is no widely accepted threshold when the number of clusters is "small," the work of Cameron, Gelbach, and Miller (2008), Angrist and Pischke (2009), and Harden (2011) suggests a cutoff of around 40 clusters. To check whether clustering in our cross-country sample with just 31 clusters produces misleading inferences, we use the wild cluster bootstrap procedure suggested by Cameron, Gelbach, and Miller (2008) for improved inference with few clusters (using Stata's *cgmwildboot* command for implementation). All results remain robust when employing the wild bootstrap procedure as an alternative to clustering.

possibility of omitted variables correlated with both teacher skills and student performance. Such omitted variables could include, for example, the educational attitude in a country: Societies that emphasize the importance of good education may have both teachers with high cognitive skills and parents who strongly support their children's education. Similarly, if the quality of the education system is persistent and not perfectly captured by our measure of parent cognitive skills, then student performance and teacher cognitive skills (who went through the same education system one generation earlier) might be positively correlated even if teacher cognitive skills have no real impact on student performance. On the other hand, sorting of students and teachers within schools and across schools (within countries) – which often plagues micro-level analysis of educational production – is no concern in our study because teacher cognitive skills are aggregated to the country level.

To the extent that omitted variables are not subject-specific, we can circumvent bias by focusing on just within-student variation in teacher skills across math and reading. Within-student effects of teacher cognitive skills on student performance are estimated with the following model:[27]

$$A_{isc,math} - A_{isc,read} = \lambda\left(T_{c,numeracy} - T_{c,literacy}\right) + \gamma_1\left(P_{isc,numeracy} - P_{isc,literacy}\right) + \left(\boldsymbol{I}_{sc,math} - \boldsymbol{I}_{sc,read}\right)\gamma_2 + \left(\varepsilon_{isc,math} - \varepsilon_{isc,read}\right).$$

This model holds constant all factors that do not differ between subjects, capturing subject-invariant performance differences across students (e.g., family background, innate ability, and motivation) and across countries (e.g., general educational attitude).[28]

The within-student approach, however, cannot control for unobserved differences across countries that are subject-specific. For example, if societies have both teachers with high numeracy skills and a strong preference for advancing children in math (with parents supporting their children accordingly), then fixed-effects estimates of teacher cognitive skills will still be biased. However, if a country has subject-specific preferences (e.g., for math), then the numeracy skills should be relatively high for all adults. We therefore control for the cognitive-skill levels in occupations other than teaching (e.g., managers, scientists and engineers, health professionals, business professionals) as proxies for the subject-specific preferences of countries. Alternatively, we also use as controls the average cognitive-skill level of all adults in a country and the difference in country-level adult skills between numeracy and literacy. These models exploit within-country variation in cognitive skills, placing teachers in the country-specific distribution of skills instead of relying solely on cross-country variation in teacher numeracy or literacy skills.

---

[27] Within-student across-subject variation has frequently been used in previous research (e.g., Dee (2005, (2007), Clotfelter, Ladd, and Vigdor (2010), and Lavy (2015)).

[28] In contrast to the OLS estimates, the estimated effect of teacher cognitive skills in the student fixed-effects model is "net" of teacher skill spillovers across subjects (for example, if teacher literacy skills affect student math performance). Spillover effects are completely eliminated when cross-subject spillovers are identical in math and reading.

5. Teacher Cognitive Skills and Student Performance

It is easiest to motivate the analysis with simple visual evidence showing that teacher cognitive skills are positively associated with student performance aggregated to the country level. The two upper graphs in Figure 3 show the unconditional cross-country relationship between teacher numeracy skills and student math performance (left panel) and between teacher literacy skills and student reading performance (right panel). Both numeracy and literacy skills of teachers are clearly positively associated with aggregate student performance. The two bottom graphs in Figure 3 show specifications analogous to those in the upper panel but additionally include country-specific skills of all adults aged 25–65 to net out the skill persistence across generations.[29] The coefficient on teacher numeracy skills is reduced only modestly, while the coefficient on teacher literacy skills even increases.[30] Although coarse, these country-level plots indicate that teacher cognitive skills could be a determinant of international differences in student performance.[31]

5.1 Ordinary Least Squares Results

We now more rigorously investigate the relationship between teacher cognitive skills and student performance using student-level test-score data. Table 2 reports results from the least squares estimation of Equation (1). The unconditional correlation between teacher numeracy skills and individual-level student math performance (Column 1) is identical to the country-level estimate presented in Figure 3. The coefficient on teacher numeracy skills remains significant when adding a large set of background factors at the individual, family, school, and country level (Column 2) and when including the numeracy skills of parents of PISA students (Column 3). The estimate in Column 3 implies that a one SD increase in teacher numeracy skills increases student math performance by almost 0.15 SD. Even though various parent characteristics, such as education level and number of books at home, are included, parent numeracy skills are significantly related to student performance, but the coefficient is rather modest in size compared to teacher cognitive skills. Columns 4–6 report results for reading. The coefficient on teacher literacy skills is highly statistically significant across specifications, although it is somewhat smaller than the coefficient on teacher numeracy skills in the specification with all controls (0.09; see Column 6).[32]

---

[29] The country-level correlations between teacher skills and adult skills are 0.77 for numeracy and 0.86 for literacy. Skills of teachers and adults are substantially correlated since both have been educated in the same education system at about the same time.

[30] When omitting teacher skills, adult skills and student performance are strongly positively correlated in both math and reading. However, when conditioning on teacher skills, the estimates for adult skills substantially decrease in size and lose statistical significance.

[31] Note that Korea and Singapore are outliers in the plots that control for average adult skills. This likely reflects the dramatic changes in educational performance of these societies over time (OECD (2016)).

[32] Using average teacher skills instead of median teacher skills leads to very similar results, with a coefficient on numeracy (literacy) skills of 0.138 (0.093). Both coefficients are statistically significant at the 1 percent level.

When accounting for student and family influences on student performance (see Columns 2 and 5), the impact of teacher skills decreases considerably more in reading than in math. This suggests that parents are more important for improving their children's reading abilities than their math performance. These results, and the smaller impact of teacher literacy skills compared to teacher numeracy skills, are consistent with the common finding that students' reading scores are more difficult to improve by teachers than their math scores.

The estimated coefficients on the other control variables included in Columns 3 and 6 are reported in Appendix Table A-6. All coefficients have the expected signs. For example, girls perform worse in math than boys but perform better in reading; and migrants perform worse than natives in both subjects. Student performance is positively associated with the number of books at home (a proxy for the educational, social, and economic background), parents' education degree, and the skill content of parents' occupation. Students perform better in urban schools, in private schools, in schools with less shortage of teachers, and in schools with more subject-specific instruction time (only significant for math). Regarding the country-level characteristics, we observe a zero coefficient on educational expenditure per student, while school starting age is positively related to student performance.

We also find some evidence for heterogeneity of the teacher-skill effect across student subgroups (Table A-7). The impact is somewhat larger for girls than for boys, for low-SES students compared to high-SES students[33] (particularly in reading), and for natives relative to migrants (particularly in math).[34] Parent cognitive skills are considerably more important for high-SES students (with a significantly positive coefficient also for literacy), while there are no differences by gender or migration status.

The results indicate that students living in the countries at the top of the PISA rankings perform better in math and reading in part because their teachers have higher numeracy and literacy skills. To gauge the magnitude of our estimates, we use these OLS coefficients to simulate the improved student performance if each country brought its teachers up to the level of Finnish teachers, who are the most skilled teachers by the PIAAC measures (Table 3). For some countries, such as Japan, this is not a huge change, but even Japanese students would improve somewhat (0.06 SD in math and 0.02 SD in reading). But for other countries, the improvements in student performance would be substantial. The U.S. would be expected to improve by roughly 0.33 SD in math; Turkey and Chile, being at the

[33] SES status is measured by the PISA index of economic, social, and cultural status (ESCS).

[34] Because first-generation migrants might have migrated to the PISA test country shortly before the PISA test, we cannot ascribe their math and reading performance to the skill level of teachers in the test country. Therefore, we use only second-generation migrants in this analysis since these students were born in the PISA test country and have spent their school career in the education system of that country.

bottom of the international league table, would be expected to improve by about 0.54 SD and 0.57 SD, respectively, in math.

The teacher-skill estimates do not capture the effect of just a single school year but rather reflect the cumulative effect of teacher cognitive skills on student performance over all school years. Thus, these are long-run impacts that presume that the quality of students' teachers in the first ten grades would improve to the level of Finland – something that would take time and effort to realize.

The baseline OLS model already controls for a multitude of determinants of student performance, including a proxy for the cognitive skills of parents. Still, as teacher skills are measured at the country level, identification also raises particular challenges in this international setting. While we control for cross-country differences in educational expenditures and school starting age, countries may also differ from one another in other, hard-to-observe ways. For instance, cultural traits, educational attitudes, and the nature of teacher preparation may be associated with both teacher cognitive skills and student performance. To circumvent potential biases due to unobserved country heterogeneity that is similar across math and reading, we employ a student fixed-effects model (Section 5.3). Additionally, we exploit variation in cognitive skills within countries to address international differences in subject-specific confounding factors (Section 5.4). We also provide a simple placebo test that addresses the possibility that teacher skills are proxies for overall country-specific skill patterns (Section 5.5). Before presenting these estimates, however, we first show that the baseline OLS results are robust to controlling for country-specific measures of instructional practices and educational institutions.

5.2 Robustness Checks

One worry is that our subject-specific teacher-skill measures are confounded by correlated differences in pedagogical skills. To investigate this, we use information from the PISA students about their teachers' activities in language and math classes to construct indicators of subject-specific instructional activities as proxies for teachers' pedagogical skills. We follow the OECD (2010) approach of measuring specific instructional practices through survey responses of students (e.g., how often does a teacher ask questions that make students reflect on a problem), while we aggregate these instructional practices to the school level.[35] As noted in Section 3, instructional practices are asked

[35] For *reading*, we use the following items (each measured on a 4-point scale ranging from "never or hardly ever" to "in all lessons") to construct the instructional-practice indicator: asking students to explain the meaning of a text; asking questions that challenge students to get a better understanding of a text; giving students enough time to think about their answers; recommending books or author to read; encouraging students to express their opinion about a text; helping students relate the stories they read to their lives; and showing students how the information in texts builds on what they already know. For *math*, we use the following items (each measured on a very similar 4-point scale ranging from "never or rarely" to "almost or almost always"): asking questions that make students reflect on the problem; giving problems that require students to think for an extended time; presenting problems in different contexts so that students know whether they have understood the concepts; helping students to learn from mistakes they have made; asking students to explain

only for the subject that was the focus in the respective PISA cycle (reading in 2009 and math in 2012). For the PISA cycle when a subject was not the focus, we impute the subject-specific instructional-practice indicator by using the country-level measure from the other PISA survey, assuming that the instructional practices in a subject have not noticeably changed within a country over the three-year period between 2009 and 2012.[36]

Table 4 reports the results when we augment the baseline model by controls for the instructional practices in math and language classes (the baseline estimates are reported in Column 1 for math and in Column 3 for reading). The instructional-practice indicators are positively related to student performance, but are statistically insignificant. Importantly, however, when instructional practices are added, the teacher-skill estimates change very little, suggesting that teacher cognitive skills have an independent impact on student performance.[37]

Supporting this, we construct another indicator using information on instructional practices reported by teachers.[38] In line with the results in Table 4, all teacher-reported instructional practices are negatively correlated with teacher cognitive skills, suggesting that, if anything, the impacts of cognitive skills are understated by omitting the pedagogical skills of teachers.[39]

Another worry is that teacher cognitive skills are correlated with other country-level factors that affect student performance. For example, both teacher cognitive skills and student performance might be higher in more developed countries. Also, countries with higher teacher skills might also have educational institutions that are more supportive of student learning. To rule out these potential confounds, we additionally control for GDP per capita (as a measure of a country's state of development), teacher gross hourly wage (as a proxy for general teacher quality), the existence of teacher performance pay (to capture teacher-incentive and teacher-sorting mechanisms), and central exit exams (reflecting other achievement-enhancing institutions). Adding these country-level factors does not change the impact of teacher cognitive skills (Table A-8).

---

how they have solved a problem; and presenting problems that require students to apply what they have learnt to new contexts.

[36] To some extent, the country-level instructional-practice indicators just reflect cultural differences in how actively teachers communicate with their students. Therefore, it is understandable that the instructional-practice measure is largest in Anglo-Saxon countries and smallest in Asian countries.

[37] The coefficients on teacher cognitive skills even increase slightly since teacher cognitive skills and instructional practices are negatively correlated at the country level (r=-0.06 in math and r=-0.45 in reading).

[38] Data come from TALIS 2013 (see OECD (2014) for details). Instructional practices assessed in TALIS include: present a summary of recently learned content; students work in small groups to come up with a joint solution to a problem or task; give different work to the students who have difficulties learning and/or to those who can advance faster; refer to a problem from everyday life or work to demonstrate why new knowledge is useful; let students practice similar tasks until teacher knows that every student has understood the subject matter; check students' exercise books or homework; students work on projects that require at least one week to complete; students use ICT for projects or class work.

[39] We do not use instructional practices from TALIS in the student-level regressions because nine of the 31 countries in our sample (Austria, Germany, Greece, Ireland, Lithuania, New Zealand, the Russian Federation, Slovenia, and Turkey) did not participate in TALIS 2013, which would substantially reduce our sample.

Finally, to address issues of divergent national cultures (in particular, differing educational attitudes) around the world, we show that our results are robust to specifications that include continental fixed effects and that restrict the analysis to just European countries, which makes the sample culturally more homogeneous (Table A-9). Moreover, any analysis that exploits international variation with limited degrees of freedom might suffer from the problem that the results are driven by a few outliers. Therefore, we replicated the baseline OLS specification with all control variables, but sequentially excluded each country from the sample. The estimated teacher-skill effects are always very close to the baseline coefficients, confirming that the results are not driven by any individual country (results available upon request).

In summary, the estimated impact of teacher cognitive skills on student performance proves highly robust to additional controls and to using different samples.

5.3 Student Fixed-Effects Results

While the previous section has shown that our teacher-skill estimates are remarkably robust to additional controls and various subgroups, we are still concerned about omitted variables that vary at the country level. Thus, we now turn to estimation with student fixed effects. Here, we exploit only within-country variation to identify the effect of teacher cognitive skills on student performance, eliminating any non-subject-specific bias.

Table 5 presents the results of the student fixed-effects specifications that match the OLS specifications except that now the difference in student performance between math and reading is regressed on the difference in teacher skills between numeracy and literacy. Also, all control variables that differ across subjects are included in first differences. Across specifications, the student fixed-effects estimates for teacher cognitive skills remain sizeable and close to the OLS coefficients on teacher numeracy and literacy skills.

While neither parent cognitive skills nor teacher shortages are significantly related to student performance, the effect of instructional time on student performance is significant and similar to the effect size in Lavy (2015).[40]

5.4 Overall Country-Level Skill Differences

While the within-student approach captures any differences across countries that are not subject-specific, it does not control for subject-specific differences. However, if some countries systematically do better in math or reading – because of strong preferences for one particular subject, curricular differences, or other reasons – the estimated impacts of teacher cognitive skills may simply be a reflection of subject-specific country differences. In Table 6, we therefore additionally control

[40] Lavy (2015) exploits within-student between-subject variation using PISA data from 2006.

for the cognitive-skill level of all adults and all parents, respectively, in a country to account for countries' potential subject preference or other subject-related differences.[41] Adding the cognitive-skill levels of these broad groups does not change the impact of teachers' cognitive skills on student performance (Columns 1 and 2 for math and Columns 5 and 6 reading).

Another way to proxy for the subject-specific preferences of a country is to consider the difference between numeracy and literacy of the adult population. If the country has a strong preference for math, then it should have students with high math performance and adults whose numeracy skills are high relative to their literacy skills (vice versa for reading). However, controlling for the cognitive-skill differences of parents or adults does not affect the impact of teacher skills (Columns 3 and 4 for math and Columns 7 and 8 for reading). Interestingly, while the coefficients on these controls are statistically insignificant, they even have counterintuitive positive signs in reading.

5.5 A Simple Placebo Test

Our estimates emphasize the relationship between teacher cognitive skills and student performance, which in turn reflects the allocation of overall country talent to teaching. But, is it the skills of the teachers themselves that is affecting student performance? If the estimates do not reflect the impact of teachers per se but instead just reflect the overall skills of the society, it would be the case that relating the cognitive-skill levels of workers in occupations other than teaching could also equally explain the pattern of student achievement. To investigate this possibility, we consider a simple placebo test where we replace teacher cognitive skills with the cognitive skills of workers in 14 different occupational groups (e.g., managers, scientists and engineers, health professionals, business professionals, clerks, sales workers, service workers).[42] Of course, since all adults went through the same education system and reflect to some extent the overall skill level of the country, the cognitive skills of workers in these occupations should be positively correlated with the cognitive skills of teachers. Thus, it is hardly surprising that student performance is positively related to the cognitive-skill levels in some of these occupations (Panel A in Tables 7 and 8). Yet, especially for literacy skills, this is true only for remarkably few occupations.[43]

Most importantly, controlling for the cognitive-skill levels of workers in other occupations does not change the estimated impact of teacher cognitive skills on student performance (see Panel B in Tables 7 and 8). In contrast, the cognitive-skill levels of only few occupations are still significantly

[41] The country-specific adult skills are measured by the median skill level of all adults aged 25–65. The country-level parent cognitive skills are measured by the median skills of all PIAAC respondents aged 35–59 with children (i.e., the same PIAAC respondents used to construct the individual-level parent skills).

[42] Occupational information comes from PIAAC and refers to respondents' current job. The 14 occupations considered here cover the full range of a country's occupational distribution.

[43] Note that parent cognitive skills are more strongly related to student performance when including the cognitive-skill levels in occupations other than teaching instead of teacher cognitive skills.

positively related to student performance when teacher skills are included; in fact, this is true for just 3 out of 14 occupations in numeracy and not a single occupation in literacy. The impact of teacher cognitive skills remains highly significant even when we control for the cognitive-skill levels of *all* other occupations simultaneously (not shown). In this highly demanding specification, almost none of the cognitive-skill levels of other occupations are significantly related to student performance (except numeracy skills of craft workers and literacy skills of elementary workers). Thus, it is the allocation of skills to teaching and not a country's overall skill level that appears to matter for student performance.

6. Determinants of Teacher Cognitive Skills

The existing international differences in teacher cognitive skills reflect both where teachers are drawn from in each country's skill distribution and the overall level of cognitive skills in each country's population – and policies to improve the skills of teachers could conceptually focus on either of these dimensions. Increasing the overall achievement of a country's population would of course be highly desirable and would be self-reinforcing through improving the pool of potential teachers. Nonetheless, consideration of potential overall improvement policies, while widely discussed elsewhere, is beyond the scope of this analysis.

We instead focus on the determinants of where teachers are drawn from the overall skill distribution of the population, which as noted above has received relatively little and narrow attention. Our international data permit a much broader investigation of how external forces and policy choices affect the skills of the teaching force. Specifically, we can explore across the broad range of international experiences how improvements in alternative job opportunities for women over time and differences in relative teacher pay have altered the skill levels of teachers.

6.1 Alternative Professional Opportunities for Women

Changes in the cognitive skills of teachers have been studied in the U.S., where there is general agreement of a decline over time in measured achievement and in other quality indicators (Murnane et al. (1991), Corcoran, Evans, and Schwab (2004a, (2004b), Bacolod (2007)).[44] A common hypothesis is that this decline in teacher cognitive skills in the U.S. during the past decades was the result of improving alternative opportunities for women in the labor market. As more women have access to high-skill, high-wage occupations, fewer high-skilled women choose to become teachers,

---

[44] There is a longer investigation of the teaching profession, largely from a sociological perspective, that focuses on the well-being of teachers in terms of their relative status and earnings, as opposed to any aspect of teacher quality or teacher effectiveness. See, for example, Bergmann (1974), Reskin (1984), and Tienda, Smith, and Ortiz (1987). Such analyses have also had an international comparative component as in Charles (1992), Blackburn, Jarman, and Brooks (2000), and Kelleher (2011), but again lacking any attention to the impact on students.

thus leading to declining average teacher skills.[45] Testing this hypothesis has been difficult, however, because the underlying data on teachers have come from piecing together a limited number of snapshots of skill differences from U.S. surveys conducted at different points in time. The limited observations plus incomplete measures of skill demands or rewards in alternative occupations present serious challenges to any analysis.[46]

In the spirit of Bacolod (2007), we relate within-country changes in labor-market choices of females to changes in teacher cognitive skills across birth cohorts. However, our analysis differs in two key ways. First, we explicitly consider the human capital intensity of alternative employment opportunities (instead of simply relying on relative average wages in teaching and elsewhere). Second, we observe multiple countries, which not only dramatically expands the range of observations but also allows us to account for any general (i.e., non-country-specific) time trends that affect both the nature of female labor-market participation and teacher skills. For example, the teaching profession might have become less attractive relative to other (possibly newly emerging) high-skilled occupations over time, explaining both an increasing share of females in other high-skilled occupations and a decline in average teacher skills.

We start by constructing an indicator of women's access to high-skilled occupations in a country's labor market. For country-cohort cells, we compute the proportion of female teachers relative to females in high-skilled occupations. We use the PIAAC micro data to classify occupations as "high-skilled" by identifying country-specific occupations that employ the most educated workers. For two-digit occupations in each country, we calculate the average years of schooling of employees currently working in each occupation at the time of the PIAAC assessment (i.e., in 2011/2012 for Round 1 and 2014/2015 for Round 2).[47] Second, ranking occupations in each country by average schooling level and starting with the occupation with the highest level, we define all occupations as "high-skilled" until males working in these occupations comprise 25 percent of all working males in the country.[48] The 25-percent rule ensures that a similar share of workers is employed in high-skilled occupations in each country; other variants of defining high-skilled occupations led to very uneven shares of males working in high-skilled occupations across countries. To obtain cohorts with sufficient numbers of teachers, we merge 15 adjacent age cohorts. As the PIAAC data cover 45 birth

[45] As Bacolod (2007) points out, the opening of alternative high-wage jobs does not necessarily imply declining teacher quality; in a Roy model, it would depend on comparative advantage in different occupations and the correlation of a worker's skills in different occupations.

[46] Bacolod (2007) expands on the data by using observations for the separate U.S. states.

[47] There are no internationally comparable data that would allow computing these country-by-cohort-specific shares on the basis of historical labor-market records.

[48] Note that teaching is a high-skilled occupation in every country in our sample. Applying an alternative categorization that classifies all occupations contained in the one-digit ISCO codes 1 (Managers) and 2 (Professionals) as high-skilled leads to qualitatively similar results.

years (excluding very young workers who mainly have not completed their university degree), we obtain three age cohorts per country.

Consistent with the notion that teacher skills are directly affected by competition from other occupations that demand high skills, we expect that higher concentrations of females in teaching lead to higher cognitive skills of teachers. The test of this exploits changes in the share of female teachers relative to women in all high-skilled occupations over three age cohorts in 23 countries.[49] Our estimation always includes cohort fixed effects to control for general time trends in women's labor-market opportunities and for skill depreciation across cohorts. Moreover, country fixed effects account for cross-country differences in women's labor-market participation and in average skill levels that are constant across birth cohorts.

Teaching remains a female-dominated profession across our sample of countries. An average of 69 percent of teachers in the 23 countries for this analysis are female. This ranges from 59 percent in Japan to 79 percent in Austria. Thus, the changing composition of female teachers noticeably affects the teaching force.

Table 9 reports the results of estimating the effect of alternative job opportunities on the skill level of teachers.[50] For both numeracy and literacy, we find that a higher share of high-skilled female workers in teaching is positively and statistically significantly related to teacher cognitive skills. The coefficients barely change when we add the average skill level of university graduates in the respective country-cohort cell to account for country-specific skill depreciation.[51]

The estimates are also economically meaningful. An increase in the share of high-skilled female workers in teaching by 10 percentage points leads to a 0.36 SD increase in the numeracy skills of teachers. (The results are slightly weaker for literacy.) The share of high-skilled female workers in teaching varies between 17 percent in Chile (18 percent in the U.S.) and 38 percent in Singapore (across all three cohorts). Thus, if females in the U.S. had similar employment opportunities as in Singapore, average teacher numeracy skills in the U.S. would increase by about 0.72 SD, bringing U.S. teachers to just above the international average in teacher numeracy skills. Across all 23 countries in the sample, the share of high-skilled female workers in teaching decreases from 29 percent in the oldest age cohort (born 1946–1960) to 22 percent in the youngest cohort (born 1976–

[49] For this analysis, we exclude the ex-communist countries (the Czech Republic, Estonia, Lithuania, Poland, Russia, the Slovak Republic, and Slovenia) and Turkey since occupational choices in these countries were less driven by market incentives but rather depended on political attitudes. While our results indicate that females' labor-market opportunities affect the level of teacher cognitive skills, the analysis uses only pseudo cohorts based on the cross-sectional PIAAC data. Thus, the validity of our results depends on the assumption that women do not change the type of their occupation (high-skilled vs. low-skilled; teacher vs. nonteacher) in a systematic way over their careers. Furthermore, our approach assumes that the country-specific pattern of skill depreciation across cohorts is similar for teachers and university graduates.

[50] Results are qualitatively similar when we use the skill level of female teachers as dependent variable.

[51] Several studies suggest that losses of skills over the life cycle occur, underlining the importance of controlling for skill depreciation (e.g., Cascio, Clark, and Gordon (2008); Edin and Gustavsson (2008)).

1990), reflecting an international improvement of alternative job opportunities for women across cohorts. This is associated with a decline of 0.25 SD in teacher numeracy skills.[52]

6.2 *Teacher Pay*

An obvious consideration in looking at the pattern of teacher skills is the pay received by teachers. In fact, the argument that teacher pay is significantly related to teacher quality has been in the heart of much of the debate about educational policy for many years (see, e.g., Dolton and Marcenaro-Gutierrez (2011)). The idea is that countries that pay teachers relatively better are able to recruit teachers from higher up in the skill distribution and also are able to retain teachers in their profession.[53] If this link is present, there would be leverage for policymakers to raise the skills of teachers in the country by paying them higher wages, with commensurate positive effects on student performance.[54]

In order to investigate the salary-skills relationship across countries, we first estimate whether *ceteris paribus* teachers are paid a premium in the labor market. Using the individual-level PIAAC data, we estimate a Mincer-like earnings equation with log earnings (ln y) regressed on gender (G), potential work experience (E), achievement in numeracy and literacy (A), and a teacher indicator (T).[55]

$$\ln y = a_0 + a_1 G + a_2 E + a_3 E^2 + \boldsymbol{A a}_4 + \delta T + \varepsilon \qquad (2)$$

The coefficient δ is the premium for teachers given their characteristics. We estimate a separate premium for each country, and we find a wide dispersion. Figure 4 shows the estimated teacher premiums across countries, ranging from +45 percent in Ireland to -22 percent in the United States

[52] Note that skill depreciation over the lifecycle is accounted for by including both cohort fixed effects and the cohort-specific skill level of nonteacher college graduates.

[53] Raising pay might also provide already-recruited teachers with more incentives to exert higher effort to improve the educational outcomes of the children they teach. The evidence on effort is, however, not very encouraging; see Springer et al. (2010). While much of the policy discussion of performance pay does not distinguish between the effort margin and the selection-retention margin, it is the latter that seems more important. The international studies effectively look at selection and retention, while within-country analyses almost always look at effort; see Woessmann (2011). For developing countries, the evidence on effort is stronger (see Muralidharan and Sundararaman (2011)), but this might not generalize to the developed countries we analyze.

[54] Another channel through which a positive association between teacher pay and teacher skills may materialize (at least in the long run) is that higher salaries for teachers may improve the status of the teaching profession. As a result, more children might want to become teachers in the future, facilitating the recruitment of more able individuals.

[55] This approach follows Hanushek et al. (2015) in estimating an earnings function without years of schooling, which is one of several inputs into cognitive skills. We use the sample of all university graduates surveyed in PIAAC in each country, which are the relevant comparison group for teachers (88 percent of teachers have obtained a college degree). However, results are qualitatively similar when we add years of schooling as an additional control or estimate the Mincer earnings function on the whole population.

and Sweden.[56] (Table A-10 presents the detailed regression output for each country). While there have been many discussions of the relative pay of teachers in the United States (see Hanushek (2016)), most have ignored the possibility that teachers are systematically different from college graduates working in other occupations (e.g., in terms of cognitive skills and gender composition). The estimates here indicate that teachers are paid some 20 percent less than comparable college graduates elsewhere in the U.S. economy after adjusting for observable characteristics.

Table 10 puts teacher pay and teacher skills together. Here, the cognitive skills of teachers in the country are regressed on the teacher wage premiums (δ). Importantly, estimates are conditioned on the cognitive skills of all nonteacher college graduates to account for international differences in overall country skill levels and to allow us to assess how pay relates to the position of teachers in the distribution of the country's skills.[57]

The results, shown graphically in Figure 5, indicate that higher relative teacher pay is systematically related to higher teacher skills. The clear conclusion is that countries that pay teachers more for their skills also draw their teachers from higher parts of the skill distribution. In terms of magnitude, a 10 percentage points higher teacher wage premium is associated with an increase in teacher skills of about 0.10 SD.[58] The coefficient on college graduates' skills is close to 1 for both numeracy and literacy, again suggesting the powerful influence of a country's overall skill level.

These results are also consistent with previous work in the U.S. on pay-skill relationships. Corcoran, Evans, and Schwab (2004b) argue that, while average cognitive skills of teachers have not changed much, there has been a sharper decline in the top deciles of skills. Bacolod (2007) finds larger declines in teacher cognitive skills. Both see the importance of teacher salaries and alternative opportunities for women in the labor market.

The interpretation of these results is, however, important for policy. These estimates are reduced-form estimates that reflect the labor-market equilibrium. They do not, however, indicate what the supply function for higher quality teachers looks like. In other words, they are not causal estimates of how the quality of teachers would change if teacher salaries were raised.[59] Moreover, the estimated

---

[56] It is remarkable that teacher wage premiums are similarly low in the United States and Sweden, since both countries are at opposite extremes of wage inequality (see Table 1 in Hanushek et al. (2015)). In the United States, workers at the 90th percentile of the wage distribution earn 4.5 times as much as workers at the 10th percentile. In Sweden, workers at the 90th percentile earn only twice as much as workers at the 10th percentile.

[57] An alternative approach is to run country-level regressions of teacher skills on relative teacher wages, measured as the percentile rank of country-specific mean teacher wages in the wage distribution of all nonteacher college graduates. This approach yields similar salary-teacher skill results, but it does not allow for any differences in the distribution of earnings characteristics between teachers and nonteachers.

[58] These estimates are likely downward biased because the teacher wage premiums are estimated coefficients and therefore contain error. Assuming that the errors are heteroscedastic (as they come from separate regressions), the true coefficients are slightly larger by 5 percent.

[59] These issues have been part of the policy discussion in the U.S., where questions have arisen about how to attract more effective teachers as measured by teacher value-added. Higher teacher salaries would undoubtedly expand the pool

relationship relates to the long run after many cohorts of teachers have been recruited. In other words, while making it clear that a more skilled teaching force will require higher salaries, the evidence says nothing about either how salaries should be structured or the responsiveness of teachers to higher salary offers.[60]

7. A Larger Policy Perspective

Our results indicate that the overall level of cognitive skills of a country's teacher force directly influences the student achievement levels that a country can expect. The magnitudes are important, however. A one SD improvement in teacher cognitive skills leads to a 0.15 (0.09) SD improvement in PISA scores in math (reading). Since PISA scores represent the cumulative learning of 15-year-olds, this suggests an average learning gain of about 0.01-0.015 SD per year.

An appropriate comparison for policy purposes would be estimates of the variations in teacher effectiveness derived from value-added models of annual learning gains (Hanushek and Rivkin (2012)). A consensus estimate is that a one SD difference in teacher effectiveness (measured by value-added) relates to a student performance difference of around 0.15 SD per year.[61] These results suggest that a pool of teachers with higher cognitive skills tends to contain more effective teachers, but within that pool there is a wide variation in teacher effectiveness. Thus, knowing a potential teacher's cognitive skills provides some overall indication of potential effectiveness but at the individual teacher level other factors exert greater influence on effectiveness.[62]

We currently have virtually no information about either the supply function of teacher cognitive skills or the supply function of teacher effectiveness. The previous analysis suggests, not surprisingly, that average teacher cognitive skills move with pay premiums in teaching. There is also some

---

of potential teachers and would also help to cut down on teacher turnover. This evidence does not, however, indicate that more effective teachers will be hired out of the enlarged pool; nor does it indicate that the teachers who are induced to stay in teaching are the more effective teachers. The same holds for changing the cognitive skills of the teaching force.

[60] In a separate analysis, we have investigated whether relative public-sector wages (i.e., mean public-sector wages over mean private-sector wages) affects an individual's decision to enter the teaching profession. Using annual OECD data on public-sector and private-sector wages for multiple countries, we aggregated the data to the same three birth cohorts as in Section 6.1. Controlling for country and birth-cohort fixed effects, we fail to find a robust relationship between teacher cognitive skills and the relative public-sector wages in the years before college graduation. There are several potential reasons for this result. Most importantly, we do not observe teacher wages, but rather rely on coarse measures of average public and private wages. Furthermore, it is unclear at which point in their educational career individuals decide to become teachers. We also made a preliminary investigation of considering economic conditions at the beginning of careers on teacher skills (following Nagler, Piopiunik, and West (2015)), but the small samples when finely disaggregated by age could not support this estimation.

[61] This is a conservative estimate since it is based on just the within-school variance in teacher quality. In terms of total variance that includes any between-school variation in quality, one SD higher teacher effectiveness is related to about 0.2-0.3 SD better student performance.

[62] This interpretation does help to reconcile previous findings about the inconsistent impact of cognitive skills in determining individual teacher effectiveness (Hanushek (2003)). This evidence, all drawn from U.S. experiences, is consistent with individual score test differences being hard to disentangle from variations in overall effectiveness, implying that large samples with good measures of cognitive skills are necessary in order to detect the impacts.

indication that teacher effectiveness responds to pay differentials through exit and retention decisions of teachers.[63] These do not provide sufficient information to derive supply elasticities on either margin, making it impossible to compare directly the achievement implications of alternative pay programs. While this analysis has focused on the achievement effects of altering the average level of cognitive skills of teachers, it is unlikely that the appropriate policy would be simply moving the entire teacher distribution up as opposed to altering the shape of the distribution along with any increase in average salaries (Hanushek (2016)). Nonetheless, how skills and effectiveness might, for example, respond to removing the large discount to teacher salaries in the United States (see Figure 4) is unknown.

8. Conclusions

We use newly available data from the Programme for the International Assessment of Adult Competencies (PIAAC) to provide a novel description of the skills of teachers in numeracy and literacy in 31 developed economies. These teacher cognitive skills differ substantially across countries. We then combine the country-level measures of teacher cognitive skills with micro data on student performance from PISA to estimate international education production functions with rich controls for student, school, and country background factors, including coarse measures of the cognitive skills of the parents of PISA students. In addition to OLS models, we estimate the impact of teacher cognitive skills using student fixed-effects models, which exploit between-subject variation and account for omitted (non-subject-specific) country-level factors.

With both approaches, we find a consistent relationship between teacher cognitive skills and student performance. In terms of magnitude, a one SD increase in teacher cognitive skills is associated with an increase in student performance at age 15 of about 0.15 SD in math and 0.09 SD in reading. Alternative specifications that control for the cognitive skills of all adults in a country or of workers in occupations other than teaching indicate that the teacher-skill effects are not simply reflecting overall differences in skills among countries but instead are directly related to where teachers are drawn from in the country's skill distribution.

We then consider possible determinants of teacher cognitive skills. Exploiting within-country changes in the share of women working in high-skilled occupations outside teaching across three birth cohorts in 23 countries, we find that a larger share of women in high-skilled jobs other than teaching is significantly related to a lower cognitive-skill level of teachers. Differences in women's access to high-skilled occupations represent one determinant of the observed international differences

[63] Dee and Wyckoff (2015) investigate the strong effectiveness-related salary system in Washington, DC. Large potential pay incentives encourage highly effective teachers to do even better; a threat of firing poor performers encourages increased exit from the system at the bottom end of performance.

in teacher cognitive skills and of the time pattern of changing teacher skills. We also show that wage premiums paid to teachers (given their gender, work experience, and cognitive skills) are directly related to teacher cognitive skills in a country. Note though that the key is whether teachers are paid a premium, because wages in all occupations, not just teaching, respond to higher cognitive skills (Hanushek et al. (2013, (2015)).

Within-country evidence, primarily from the United States, has highlighted the importance of teacher quality for student achievement. But the research behind this has been largely unable to identify any characteristics or behavior of teachers that systematically lead to higher effectiveness. By considering international differences in student performance, the analysis here is able to identify an important role for better cognitive skill of teachers as an ingredient into teacher effectiveness. Simply put: Smarter teachers produce smarter students.

References

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Auguste, Byron, Paul Kihn, and Matt Miller. 2010. *Closing the talent gap: Attracting and retaining top-third graduates to careers in teaching*: McKinsey and Company (September).

Bacolod, Marigee P. 2007. "Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Quality." *Review of Economics and Statistics* 89, no. 4 (November): 737-751.

Barber, Michael, and Mona Mourshed. 2007. *How the world's best-performing school systems come out on top*: McKinsey and Company.

Bergmann, Barbara R. 1974. "Occupational Segregation, Wages and Profits When Employers Discriminate by Race or Sex." *Eastern Economic Journal* 1, no. 2 (April): 103-110.

Björklund, Anders, and Kjell G. Salvanes. 2011. "Education and family background: Mechanisms and policies." In *Handbook of the Economics of Education, Vol. 3*, edited by Stephen Machin Eric A. Hanushek and Ludger Woessmann. Amsterdam: North Holland: 201-247.

Blackburn, Robert M., Jennifer Jarman, and Bradley Brooks. 2000. "The Puzzle of Gender Segregation and Inequality: A Cross-National Analysis." *European Sociological Review* 16, no. 2: 119-135.

Blossfeld, Hans-Peter, Hans-Günther Roßbach, and Jutta von Maurice, eds. 2011. *Education as a lifelong process: The German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft*. Special Issue 14-2011.

Britton, Jack, and Carol Propper. 2016. "Teacher pay and school productivity: Exploiting wage regulation." *Journal of Public Economics* 133(January): 75-89.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90, no. 3 (August): 414-427.

Cascio, Elizabeth, Damon Clark, and Nora Gordon. 2008. "Education and the Age Profile of Literacy into Adulthood." *The Journal of Economic Perspectives* 22, no. 3 (Summer): 47-70.

Charles, Maria. 1992. "Cross-National Variation in Occupational Sex Segregation." *American Sociological Review* 57, no. 4 (August): 483-502.

Chetty, Raj, John N. Friedman, and Jonah Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review* 104, no. 9 (September): 2633–2679.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources* 45, no. 3 (June 20, 2010): 655-681.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

Corcoran, Sean P., William N. Evans, and Robert M. Schwab. 2004a. "Changing labor-market opportunities for women and the quality of teachers, 1957-2000." *American Economic Review* 94, no. 2 (May): 230-235.

Corcoran, Sean P., William N. Evans, and Robert M. Schwab. 2004b. "Women, the labor market, and the declining relative quality of teachers." *Journal of Policy Analysis and Management* 23, no. 3 (Summer): 449-470.

Dee, Thomas S. 2005. "A teacher like me: Does race, ethnicity, or gender matter?" *American Economic Review* 95, no. 2: 158-165.

Dee, Thomas S. 2007. "Teachers and the gender gaps in student achievement." *Journal of Human Resources* 42, no. 3 (Summer): 528-554.

Dee, Thomas S., and James Wyckoff. 2015. "Incentives, selection, and teacher performance: Evidence from IMPACT." *Journal of Policy Analysis and Management* 34, no. 2 (Spring): 267-297.

Dolton, Peter J., and Wilbert van der Klaauw. 1999. "The turnover of teachers: A competing risks explanation." *Review of Economics and Statistics* 81, no. 3 (August): 543-552.
Dolton, Peter, and Oscar D. Marcenaro-Gutierrez. 2011. "If you pay peanuts do you get monkeys? A cross-country analysis of teacher pay and pupil performance." *Economic Policy* 26, no. 65 (January): 5-55.
Donald, Stephen G., and Kevin Lang. 2007. "Inference with difference-in-differences and other panel data." *Review of Economics and Statistics* 89, no. 2: 221-233.
Edin, Per-Anders, and Magnus Gustavsson. 2008. "Time Out of Work and Skill Depreciation." *Industrial & Labor Relations Review* 61, no. 2 (January): 163-180.
Eide, Eric, Dan Goldhaber, and Dominic Brewer. 2004. "The teacher labour market and teacher quality." *Oxford Review of Economic Policy* 20, no. 2: 230-244.
Glewwe, Paul, Eric A. Hanushek, Sarah D. Humpage, and Renato Ravina. 2013. "School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010." In *Education Policy in Developing Countries*, edited by Paul Glewwe. Chicago: University of Chicago Press: 13-64.
Grönqvist, Erik, and Jonas Vlachos. 2016. "One size fits all? The effects of teachers' cognitive and social abilities on student achievement." *Labour Economics* 42(October): 138-150.
Hanushek, Eric A. 2002. "Publicly provided education." In *Handbook of Public Economics, Vol. 4*, edited by Alan J. Auerbach and Martin Feldstein. Amsterdam: North Holland: 2045-2141.
Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *Economic Journal* 113, no. 485 (February): F64-F98.
Hanushek, Eric A. 2006. "School resources." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 865-908.
Hanushek, Eric A. 2016. "School human capital and teacher salary policies." *Journal of Professional Capital and Community* 1, no. 1: 23-40.
Hanushek, Eric A., Susanne Link, and Ludger Woessmann. 2013. "Does school autonomy make sense everywhere? Panel estimates from PISA." *Journal of Development Economics* 104(September): 212-232.
Hanushek, Eric A., and Richard R. Pace. 1995. "Who chooses to teach (and why)?" *Economics of Education Review* 14, no. 2 (June): 101-117.
Hanushek, Eric A., and Steven G. Rivkin. 2006. "Teacher quality." In *Handbook of the Economics of Education, Vol. 2*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 1051-1078.
Hanushek, Eric A., and Steven G. Rivkin. 2010. "Generalizations about using value-added measures of teacher quality." *American Economic Review* 100, no. 2 (May): 267-271.
Hanushek, Eric A., and Steven G. Rivkin. 2012. "The distribution of teacher quality and implications for policy." *Annual Review of Economics* 4: 131-157.
Hanushek, Eric A., Guido Schwerdt, Simon Wiederhold, and Ludger Woessmann. 2013. "Returns to skills around the world." NBER Working Paper 19762. Cambridge, MA: National Bureau of Economic Research (December).
Hanushek, Eric A., Guido Schwerdt, Simon Wiederhold, and Ludger Woessmann. 2015. "Returns to skills around the world: Evidence from PIAAC." *European Economic Review* 73: 103-130.
Hanushek, Eric A., and Ludger Woessmann. 2011. "The economics of international differences in educational achievement." In *Handbook of the Economics of Education, Vol. 3*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann. Amsterdam: North Holland: 89-200.
Hanushek, Eric A., and Ludger Woessmann. 2015. *The knowledge capital of nations: Education and the economics of growth*. Cambridge, MA: MIT Press.
Harden, Jeffrey J. 2011. "A Bootstrap Method for Conducting Statistical Inference with Clustered Data." *State Politics & Policy Quarterly* 11, no. 2 (June): 223-246.
Imbens, Guido W., and Michal Kolesar. 2012. "Robust Standard Errors in Small Samples: Some Practical Advice." NBER Working Paper No. 18478. Cambridge, MA: National Bureau of Economic Research (October).

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2008. "What does certification tell us about teacher effectiveness? Evidence from New York City." *Economics of Education Review* 27, no. 6 (December): 615-631.

Kelleher, Fatimah. 2011. *Women and the teaching profession: Exploring the feminisation debate*. London and Paris: Commonwealth Secretariat and UNESCO.

Kershaw, Joseph A., and Roland N. McKean. 1962. *Teacher shortages and salary schedules*. NY: McGraw-Hill.

Lavy, Victor. 2015. "Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries." *Economic Journal* 125, no. 588: F397-F424.

Leigh, Andrew. 2013. "The Economics and Politics of Teacher Merit Pay." *CESifo Economic Studies* 59, no. 1 (March): 1-33.

Loeb, Susanna, and Marianne E Page. 2000. "Examining the link between teacher wages and student outcomes: The importance of alternative labor market opportunities and non-pecuniary variation." *Review of Economics and Statistics* 82, no. 3 (August): 393-408.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119, no. 1: 39-77.

Murnane, Richard J., Judith D. Singer, John B. Willett, James J. Kemple, and Randall J. Olsen. 1991. *Who will teach? Policies that matter*. Cambridge, MA: Harvard University Press.

Nagler, Markus, Marc Piopiunik, and Martin R. West. 2015. "Weak markets, strong teachers: recession at career start and teacher effectiveness." NBER Working Paper 21393. Cambridge, MA: National Bureau of Economic Research.

OECD. 2010. *PISA 2009 Results: What Makes a School Successful? – Resources, Policies and Practices*. Vol. IV. Paris: Organisation for Economic Cooperation and Development.

OECD. 2013. *OECD skills outlook 2013: First results from the survey of adult skills*. Paris: Organisation for Economic Co-operation and Development.

OECD. 2014. *TALIS 2013 Results: An International Perspective on Teaching and Learning*. Paris: OECD.

OECD. 2016. *Skills matter: Further results from the survey of adult skills*. Paris: Organisation for Economic Co-operation and Development.

Reskin, Barbara F., ed. 1984. *Sex Segregation in the Workplace: Trends, Explanations, Remedies*. Washington, DC: National Academy Press.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73, no. 2 (March): 417-458.

Rockoff, Jonah E. 2004. "The impact of individual teachers on student achievement: Evidence from panel data." *American Economic Review* 94, no. 2 (May): 247-252.

Schleicher, Andreas. 2013. "What teachers know and how that compares with college graduates around the world." In *Education Today: Global perspectives on education and skills*. Paris: OECD.

Springer, Matthew G., Dale Ballou, Laura Hamilton, Vi-Nhuan Le, J.R. Lockwood, Daniel F. McCaffrey, Matthew Pepper, and Brian M. Stecher. 2010. *Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching*. Nashville, TN: National Center on Performance Incentives, Vanderbilt University.

Tienda, Marta, Shelley A. Smith, and Vilma Ortiz. 1987. "Industrial Restructuring, Gender Segregation, and Sex Differences in Earnings." *American Sociological Review* 52, no. 2 (April): 195-210.

West, Martin R., and Ludger Woessmann. 2010. "'Every Catholic child in a Catholic school': Historical resistance to state schooling, contemporary private competition and student achievement across countries." *Economic Journal* 120, no. 546: F229-F255.

Woessmann, Ludger. 2003. "Schooling resources, educational institutions, and student performance: The international evidence." *Oxford Bulletin of Economics and Statistics* 65, no. 2: 117-170.

Woessmann, Ludger. 2011. "Cross-country evidence on teacher performance pay." *Economics of Education Review* 30, no. 3 (June): 404-418.

Woessmann, Ludger, Elke Luedemann, Gabriela Schuetz, and Martin R. West. 2009. *School accountability, autonomy, and choice around the world*. Cheltenham, UK: Edward Elgar.

Appendix A. Validation of PIAAC Cognitive Skills Data with External Sources

The PIAAC data on teacher cognitive skills raise two potential concerns. First, the teacher skill measures are derived from relatively small samples.Second, they rely on a new battery of achievement tests. In order to validate these measures, we compare them with estimates from larger national surveys in the United States and Germany.

We first look at the U.S. National Longitudinal Survey of Youth (NLSY79 and NLSY97). The NLSY79 is a nationally representative sample of 6,111 young men and women who were born between 1957 and 1964. The NLSY97 is a nationally representative sample of 6,748 individuals born between 1980 and 1984. (Note that these age cohorts partly overlap with the age range of the PIAAC participants.) We measure NLSY79 respondents' occupation (using four-digit Census codes) in 2010 (last available year) and NLSY97 respondents' occupation in 2011 to make this sample as comparable as possible to the PIAAC survey in 2011.[64]

We take the mathematics and language skills tested in the four AFQT subtests which are part of the Armed Services Vocational Aptitude Battery (ASVAB). The ASVAB was administered to 94 percent of NLSY79 respondents in 1980 and to 81 percent of NLYS97 respondents in 1997. We combine the scores from the mathematical knowledge and arithmetic reasoning tests into a numeracy skills measure and the scores from the word knowledge and paragraph comprehension tests into a literacy skills measure.[65] Based on these measures, teacher skills fall at the 67th (64th) percentile in the adult skill distribution in numeracy (literacy). This is quite close to the position of teacher skills in the PIAAC data for the USA (see Table 1): 70th (71st) percentile in numeracy (literacy).

We also compare teacher cognitive skills from PIAAC with those from Germany's adult cohort of the National Educational Panel Study (NEPS).[66] This dataset is a nationally representative dataset of 9,352 adults born between 1944 and 1986. NEPS has several advantages for our purpose. First, similar to PIAAC, the competency tests in NEPS aim at measuring numeracy and literacy skills in real-life situations which are relevant for labor market success and participation in society. Second, NEPS tested skills at about the same time (in 2010/2011) as PIAAC did. Third, almost the same age

---

[64] Teachers are defined as in PIAAC (i.e., excluding pre-kindergarten teachers and university professors/vocational education teachers). We weight individual-level observations with the cross-sectional weights taken from the year in which the occupation is measured, giving each NLSY survey the same total weight.

[65] As respondents were born in different years, we take out age effects by regressing test scores on year of birth dummies first (separately for NLSY79 and NYS97). We control for age effects in the NLSY data because participants were still children or adolescents at the time of testing. In contrast, we do not take out age effects in the PIAAC data because most PIAAC participants have already completed their education when tested.

[66] This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort 6 – Adults, doi:10.5157/NEPS:SC6:3.0.1. From 2008 to 2013, NEPS data were collected as part of the Framework Programme for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, the NEPS survey is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network. See Blossfeld, Roßbach, and Maurice (2011).

cohorts were tested in NEPS and PIAAC. Similar to PIAAC, we keep all adults aged 25–65 and identify teachers based on the four-digit ISCO-88 occupation codes, where occupation is measured in 2010/2011. Teacher skills in NEPS fall at the 68th (76th) percentile among the adult skill distribution in numeracy (literacy). Again, this is similar to the respective positions of teachers in the PIAAC sample for Germany: 72th (74th) percentile in numeracy (literacy).

The similarity of teacher cognitive skills in the adult skill distribution found in PIAAC and in these nationally representative datasets with larger sample sizes supports using the PIAAC scores as measures of the teacher cognitive skills in each country.