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

Eric A. Hanushek, Marc Piopiunik, Simon Wiederhold§

March 20, 2018

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. We construct country-level measures of teacher cognitive skills using unique assessment data for 31 countries. We find substantial differences in teacher cognitive skills across countries that 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 indicate a 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-skill 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, Guido Schwerdt, and Katarina Zigova for making their international data on central exit examinations available to us. Kirabo Jackson and three anonymous referees provided a variety of insightful comments and suggestions. We are also indebted to David Deming, Thomas Dohmen, 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: Catholic University Eichstätt-Ingolstadt, ifo Institute at the University of Munich, and CESifo, simon.wiederhold@ku.de.

1. Overview

Numerous international assessments have shown that student achievement differs widely across developed countries, but the source of these differences is not well-understood. While prior analysis has identified the impact of overall institutional structures (Hanushek and Woessmann (2011)), the research has been much less successful at identifying systematic features of schools and teachers that enter into explaining these achievement differences – leaving many policy discussions open to anecdotal and ad hoc explanations. This paper investigates whether differences in cognitive skills of teachers – which arise both from overall country skill differences and from policy decisions – can help explain international differences in student performance across developed countries.

Policy discussions, building largely on within-country analyses of the importance of teachers, have emphasized the role of teacher skills in 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) that for the first time allow 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 in our sample (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.

Employing a variety of estimation approaches, we consider how teacher cognitive skills are related to student achievement. While identification of causal effects is clearly difficult in this international context, the consistency of estimated impacts across alternative approaches supports the underlying importance of teacher cognitive skills. Following this basic impact estimation, complementary analyses of why teacher skills differ across countries and of the international reward structure for teacher skills provide new evidence on the sources of country differences.

We use country-level measures of subject-specific teacher skills along with rich student-level micro data from the Programme for International Student Assessment (PISA) to estimate the association of teacher cognitive skills with student performance in math and reading across 31 developed economies. The results from combining this information on teacher quality with student achievement indicate that differences in teacher cognitive skills can explain significant portions of the international differences in student performance.

Because of the obvious difficulty of reliably separating the independent impact of teacher cognitive skills from other factors potentially influencing student achievement, we pursue three different strategies to investigate the sensitivity of the estimated impacts to potential confounding factors. 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. Subject-specific parental cognitive skills, approximated with the PIAAC data, help in separating teacher impacts from the persistence of skills across generations and from smart parents.

[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.

Second, we exploit differences in the performance of students and teachers across math and reading. This student fixed-effects analysis allows us to identify the effect of teacher cognitive skills using only variation between subjects, thereby directly controlling 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, the nature of teacher labor markets, and culture.

Third, a set of alternative placebo tests strongly supports our basic estimation. First, we estimate student achievement models based on the cognitive-skill levels in other broad occupations (e.g., managers, scientists and engineers, health professionals, business professionals), but no alternative occupational grouping is systematically related to student outcomes, and estimated impacts are consistently below those of actual teachers. Second, we create pseudo-teacher samples by randomly selecting adults who match teachers on background characteristics. Estimating the student achievement models, the results again fail to rival our actual teacher estimates in terms of magnitude or significance of impact.

All empirical strategies consistently indicate a robust positive relationship between teacher cognitive skills and 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 0.10-0.15 SD higher student performance. To put these estimates into perspective, they imply that roughly one quarter of gaps in mean student performance across our 31 countries would be closed if each of these countries were to raise the median cognitive skills of teachers to the level of Finnish teachers (the most skilled teachers by the PIAAC measures).

Our results are robust to adding coarse measures of teachers' pedagogical approaches, suggesting that instructional style neither explains nor mediates the impact of teacher cognitive skills. Moreover, accounting for cross-country differences in economic development and in educational institutions such as central exit exams and controlling for continental fixed effects to

address issues of divergent national cultures do 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 United States resulting from improved alternative employment opportunities for women in the labor market during the past decades (e.g., 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-skill occupations. By observing multiple countries, we can more readily assess how female labor-market opportunities interact with teacher quality.

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

The PIAAC micro data permit looking explicitly at whether teachers in each country are 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 considerably above market and the United States and Sweden paying considerably below market. These reduced-form country-specific premiums are directly related to observed teacher cognitive skills across countries and, importantly, to student achievement differences.

Section 2 considers relevant prior research. Section 3 introduces the datasets and describes our computation of teacher cognitive skills. Section 4 presents our empirical strategies. Section 5 reports results on the association of teacher cognitive skills with 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-skill occupations and on teacher salaries. Section 7 concludes.

2. Related Literature

Large numbers of studies investigate the determinants of student achievement within individual countries.[2] This literature consistently finds 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. A parallel literature on international differences in achievement builds on the comparative outcome data in existing international student assessments. One of the clearest explanatory factors from these international studies has been the importance of family. In contrast, specific conclusions about the impact of school resources have been much more limited.

The most convincing within-country studies of the role of schools focus on differences in learning gains among teachers and classrooms. These studies of teacher value-added to student reading and math performance consistently find huge variations in teacher value-added that far exceed the impact of any measured school inputs (Hanushek and Rivkin (2012), Jackson, Rockoff, and Staiger (2014)).[3] 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 States. Second, they have not reliably described underlying determinants of teacher value-added – and in particular any determinants that can be consistently measured across countries.

Importantly, a wide range of international within- and across-country studies have generally shown that the most 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

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

[3] 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.

studies.[4]

However, two interesting contrasts coming from studies of teacher cognitive skills and of teacher salaries motivate our subsequent analysis of student achievement. Prior studies of measured teacher cognitive skills, largely from within the United States,[5] provide some suggestive results of positive impacts on student achievement.[6] However, studies incorporating measures of teacher cognitive skills have generally relied on small and idiosyncratic data sets, and the results have not been entirely consistent.[7] Nonetheless, compared to the various alternative measures of teacher quality commonly investigated, teacher test scores have been most consistently related to student outcomes. Moreover, recent work that links teacher test performance to the overall value-added of teachers also shows a positive relationship (Jackson and Bruegmann (2009), Clotfelter, Ladd, and Vigdor (2010), Jackson (2012b)).[8]

The relevant evidence on teacher salaries is different. While within-country studies tend to find that salaries are not a good measure of differences in teacher effectiveness,[9] the limited cross-country studies that are available have found salary levels to be often positively related to country

[4] For reviews of the evidence on the impact of teacher characteristics from within-country studies, see Hanushek (1995, (2003), Glewwe et al. (2013), and Woessmann (2003). For cross-country evidence, see Hanushek and Woessmann (2011). The one exception to the general conclusions is that additional years of experience at the beginning of a career quite consistently have positive and significant impacts of student achievement.

[5] For developing countries, Harbison and Hanushek (1992), Metzler and Woessmann (2012), and Bietenbeck, Piopiunik, and Wiederhold (Forthcoming) show the relevance of teacher subject knowledge using individual-level teacher data. 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 in Sweden.

[6] Measures of teacher cognitive skills were first introduced in the Coleman Report (Coleman et al. (1966)). Studies in the U.S. finding a positive impact of teacher tests on student outcomes include early studies in Hanushek (1971, (1972, (1992) and Ferguson and Ladd (1996) and, more recently, in Clotfelter, Ladd, and Vigdor (2006) and Goldhaber (2007).

[7] See Eide, Goldhaber, and Brewer (2004); Hanushek and Rivkin (2006), and the summaries in Hanushek (1997, (2003). Among the early estimates, 37 percent are positive and statistically significant and 27 percent are positive and statistically insignificant versus 10 percent negative and statistically significant and 15 percent negative and statistically insignificant. (12 percent do not provide the sign of statistically insignificant results).

[8] In related analysis, teacher test scores help in selecting effective teachers (Rockoff et al. (2011) and in identifying good teacher matches (Jackson (2012a)).

[9] Hanushek and Rivkin (2006)) provide an overview of the within-country evidence indicating that teacher salaries are a weak measure of teacher quality. However, 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.

differences in student outcomes.[10] These divergent results suggest that the salary levels of a country may be part of a country's institutional structure with important ramifications for the quality of the overall pool of potential teachers, even if the distribution of salaries within a country is not a good index of differential teacher effectiveness. Relatedly, cross-country analysis suggest that pay incentives are related to student performance even if within-country variations in pay structure are less informative.[11] The overall suggestion of the importance of salary differences across countries leads us to explore country-level teacher wage premiums and teacher cognitive skills in Section 6.

Changes in the cognitive skills of teachers have been previously 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)).[12] Bacolod (2007) documents a clear decline in the quality of young women entering the teaching profession between 1960 and 1990 that she relates to falling relative teacher wages, and Corcoran, Evans, and Schwab (2004a, (2004b) show that the decline in measured teacher skills over the period was concentrated in the upper portion of the achievement distribution.[13] Both suggest that women's opportunities to enter high-skill occupations outside teaching are a determinant of the

[10] In their country-level analysis, Lee and Barro (2001) find a positive association between teacher salary levels and student achievement. Similarly, Woessmann (2005) reports a significant positive coefficient on a country-level measure of teacher salary when added to an international student-level regression. Dolton and Marcenaro-Gutierrez (2010) pool country-level data from international tests between 1995 and 2006 to show that teacher salaries – both when measured in absolute terms and relative to wages in each country – are positively associated with student achievement, even after controlling for country fixed effects. However, since salary differentials are difficult to compare internationally, the cross-country models might be biased.

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

[12] 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.

[13] A related line of research has focused on entry and exit from teaching, investigates the importance of alternative job opportunities for teacher quality. 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). 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. None of these, however, considers teacher cognitive skills, the focus of our study. An early investigation of how preparation for and entry into teaching are related to cognitive skills is found in Hanushek and Pace (1995).

skill level of teachers in a country, a hypothesis to which we return below.[14] Importantly, the analysis of varying skill levels of teachers in these studies has not been linked directly to student performance – something that we can do for an international sample.

Throughout we focus on cognitive skills for both teachers and students. While there is increased attention to noncognitive skills both in the labor market (e.g., Heckman and Kautz (2012)) and to the role of schools and teachers in producing noncognitive outcomes (e.g., West et al. (2016) and Jackson (forthcoming)), we have no way of directly incorporating noncognitive skills in our international analysis even if they might partially be reflected in our test scores.

3. International Comparative Data

This section first describes the construction of the new international measures of teacher cognitive skills and how these skills are distributed across countries (Section 3.1). We then explain how measures of parent cognitive skills are constructed (Section 3.2) and introduce the data on student performance and further control variables (Section 3.3). Electronic Appendix A provides additional information on the datasets and the construction of variables.

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 aged 16-65 years in 33 developed economies.[15]

We define teachers as all PIAAC respondents who report a current four-digit occupation code of "primary school teacher", "secondary school teacher", or "other teacher" (which includes, for example, special education teachers and language teachers).[16] We exclude university professors and

[14] 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.

[15] None of the countries participated in both rounds of PIAAC.

[16] This includes school principals who teach, but excludes other workers at school with non-teaching occupations.

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 it is unclear whether they contribute to teaching students reading and math and because the role of this teacher group depends directly on the institutional structures of individual countries. Results are, however, very similar if we include pre-kindergarten teachers in the sample.

PIAAC does not allow us to identify the subject that a teacher is teaching, so we rely on measures of numeracy and literacy skills covering all teachers tested in PIAAC in each country.[17] We focus on the country-level median of the teacher cognitive skills because the median is more robust to outliers than the mean,[18] something that is potentially relevant in smaller samples.[19] 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.[20] 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 mean scores in the full PIAAC sample are 259 points in numeracy and 268 points in

[17] The domain-specific skills of teachers provided in PIAAC differ from subject-matter knowledge in math and reading. However, the PIAAC measures are good proxies for subject-specific knowledge as indicated by the fixed effects results, the cross-subjects results, and the placebo tests using teachers' information and communication technology (ICT) skills; see Section 5.

[18] 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.

[19] Due to the limited size of our teacher samples, we focus on the effect of median teacher skills and not other moments of the distribution in the main analysis. Considering within-country distributions, however, has no qualitative impact on the results; see footnotes 34 and 36.

[20] From the 33 countries participating in PIAAC, we exclude Cyprus (which did not participate in PISA) and Indonesia (where the PIAAC survey was administered only to the population in Jakarta).

literacy. (PIAAC measures each of the skill domains on a 500-point scale.) The range of median 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 median teacher skills of 60 points even exceeds one individual-level standard deviation (50 points). Teachers in the United States (284 points) perform worse than the sample-wide 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)). 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 5.2). 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, separately for each domain, across the 31 countries to have mean zero and standard deviation one.

Figure 1 illustrates 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 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, Russia, and Israel perform at the level of vocationally-educated Canadians. Teachers in 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 higher than those 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 country's

skill distribution and where a country's overall cognitive-skill level falls in the world distribution. As most teachers have obtained a college degree (89 percent across all PIAAC countries), we expect teacher cognitive skills to fall above the country's median. 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).

Figure 2 compares teacher cognitive skills with the skills of just college graduates in a country. 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 college skill distribution in some countries (e.g., Finland, Singapore, Ireland, and Chile) and from the lower part in others (e.g., Austria, Denmark, the Slovak Republic, and Poland).

From Table 1, teachers in Chile, France, Spain, and Turkey are drawn highest up in the country distributions of adult skills in numeracy and literacy. Although Finnish teachers are drawn from a somewhat lower part of the country's overall skill distribution, they have substantially greater skills than in Chile, France, Spain, and Turkey, reflecting the higher overall cognitive skill level in Finland. 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 (55th) percentile of the college graduate distribution in numeracy (see also Figure 2).[21]

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 Electronic Appendix B, support the overall validity of the estimates of teacher cognitive skills that are derived from PIAAC.

3.2 Parent Cognitive Skills

Because the parents of the PISA students (henceforth "PISA parents") are not tested, we use the

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

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 develop a prediction model for adult numeracy and literacy skills in this "PISA-parents sample" based on common observable characteristics that appear in both PIAAC and PISA. Specifically, separately by country, we regress the numeracy/literacy skills of PIAAC adults aged 35–59 with children (i.e., 17–44 years old when PISA students were born) on three characteristics: gender, education (3 categories), and number of books at home (6 categories).[22] The estimated coefficients from this are combined with the same three characteristics for the PISA parents in order to predict 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 family background and education in the home are important. Using parental cognitive skills adds a qualitative dimension to family influences over and above commonly employed 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 EA-1 in the Electronic Appendix 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,

[22] We compute skills separately for mothers and fathers because numeracy/literacy skills might differ across gender. 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.

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.

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.[23] 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. We use the two PISA cycles of 2009 and 2012 because the students have largely been taught by the teacher cohorts tested between 2011 and 2015 in PIAAC.[24]

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 school years between Singapore and Turkey and almost four years between Singapore and Chile. International student performance differences in reading are less pronounced but still substantial.

Summary statistics for student performance and student characteristics are reported in Table EA-2; summary statistics for parent characteristics (e.g., number of books at home and highest educational degree) in Table EA-3; summary statistics for school characteristics (e.g., weekly instructional time for language classes and math classes) in Table EA-4; and summary statistics for

[23] 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 (and science) performance. Note, however, that math scores from TIMSS are strongly correlated with math scores from PISA at the country level. For a description of the PISA assessments, see OECD (2010b).

[24] There is some disconnect in the timing of the measurement of teacher cognitive skills and when the teachers who are responsible for the performance of the 15-year-old PISA students actually taught them. The disconnect likely adds measurement error in teacher cognitive skills, which leads to a downward bias in the estimated teacher effects. While the matching of PIAAC teachers to PISA students is certainly not perfect, we assume some stability in teacher skills across adjacent age cohorts. Furthermore, there is still a large overlap of teachers in PIAAC and those who taught the PISA students since only a small fraction of teachers retires during a 10-year period and gets replaced by new, young teachers.

country characteristics (e.g., cumulative educational expenditure per student) in Table EA-5.

For the econometric analysis, we standardize student test scores at the student level separately for each subject 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.

4. Estimation Strategy

If we observed the skills of the individual teachers who teach the students tested in PISA, we would estimate the following education production function:

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

where: $\epsilon_{iksc} = \mu_{ksc} + \sigma_c + \pi_{kc} + \eta_{iksc}$.

A_{iksc} denotes the test score of student *i* in subject *k* (math or reading) in school *s* in country *c*. T_{iksc} represents the cognitive skills of student *i*'s teacher in subject *k*; the parameter λ is the focus of our attention. $\boldsymbol{F}_{isc}$ 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.[25] Finally, the error term, ε_{iksc}, is comprised of the following (unobserved) components: a school-subject-specific selection term due both to student selection into schools and classrooms and to administrative assignment to schools and classrooms, μ_{ksc}; general unmeasured differences between countries not captured by the included country-level control variables, σ_c; unmeasured country differences that are subject-specific, π_{kc}; and an idiosyncratic error term, η_{iksc}.

The consistency of the estimates of λ depends on the usual condition that:

[25] See Tables EA1–EA5 for country-specific descriptive statistics of student, parent, school, and country variables included in our regression model. Note that the shortage of teachers is not meant to capture differences in teacher skills but rather to reflect that classes may have to be skipped because there are simply not enough teachers in the school.

$$E\left(\varepsilon \middle| T, F, S, C, P, I\right) = 0\,.$$

This might fail due to omitted variables correlated with both student performance and teacher skills. For instance, λ would be biased upward if highly educated parents select schools or classrooms with higher-skilled teachers and also foster their children's learning in other ways. Similarly, student sorting across or within schools would lead to upward biased estimates if students with high (unobserved) academic ability are more likely to attend schools or classrooms with highly skilled teachers.

However, we measure teacher cognitive skills only at the country level, leading to the following baseline OLS model:

$$A_{iksc} = \alpha + \lambda \bar{T}_{kc} + \boldsymbol{F}_{isc}\boldsymbol{\beta}_1 + \boldsymbol{S}_{sc}\boldsymbol{\beta}_2 + \boldsymbol{C}_{c}\boldsymbol{\beta}_3 + \gamma_1 P_{iksc} + \boldsymbol{I}_{ksc}\boldsymbol{\gamma}_2 + \omega_{iksc}, \tag{2}$$

where: $\omega_{iksc} = \sigma_c + \pi_{kc} + \eta_{iksc}$.

Here, the cognitive skills of student *i*'s teacher (T_{iksc}) from Equation 1 are replaced by the median (subject-specific) teacher skills in country *c* ($\bar{T}_{kc}$). In contrast to micro-level analyses using skills of individual teachers, sorting of students and teachers across and within schools is not an issue in our setting since teacher cognitive skills are aggregated to the country level. Therefore, using aggregated teacher skills eliminates the unobserved student sorting component, μ_{ksc}, from the error term. However, the estimated coefficient on teacher skills might still be biased because of omitted country-level variables correlated with both teacher skills and student performance, such as 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 support their children's education.

To avoid bias due to omitted variables that do not vary across subjects, we exploit the fact that both students and teachers were tested in two subjects and ask whether *differences* in teacher cognitive skills between numeracy and literacy are systematically related to *differences* in student performance between math and reading. Thus, we identify the effect of teacher cognitive skills

based only on variation between teacher numeracy and literacy skills within the same student.[26] The within-student model is derived by subtracting the OLS model for reading (Equation 3) from the OLS model for math (Equation 4):

$$(3) \quad A_{isc,read} = \alpha + \lambda \bar{T}_{c,literacy} + \boldsymbol{F}_{isc}\boldsymbol{\beta}_1 + \boldsymbol{S}_{sc}\boldsymbol{\beta}_2 + \boldsymbol{C}_c\boldsymbol{\beta}_3 + \gamma_1 P_{isc,literacy} + \boldsymbol{I}_{sc,read}\boldsymbol{\gamma}_2 + \omega_{isc,read}$$

$$(4) \quad A_{isc,math} = \alpha + \lambda \bar{T}_{c,numeracy} + \boldsymbol{F}_{isc}\boldsymbol{\beta}_1 + \boldsymbol{S}_{sc}\boldsymbol{\beta}_2 + \boldsymbol{C}_c\boldsymbol{\beta}_3 + \gamma_1 P_{isc,numeracy} + \boldsymbol{I}_{sc,math}\boldsymbol{\gamma}_2 + \omega_{isc,math}.$$

This yields the following within-student across-subject model that eliminates any non-subject-specific differences across students ($\boldsymbol{F}_{isc}$), schools ($\boldsymbol{S}_{sc}$), and countries (observed factors, $\boldsymbol{C}_c$, and unobserved factors, σ_c):

$$(5) \quad A_{isc,math} - A_{isc,read} = \lambda\left(\bar{T}_{c,numeracy} - \bar{T}_{c,literacy}\right) + \gamma_1\left(P_{isc,numeracy} - P_{isc,literacy}\right) + \left(\boldsymbol{I}_{sc,math} - \boldsymbol{I}_{sc,read}\right)\gamma_2 + \left(\upsilon_{isc,math} - \upsilon_{isc,read}\right),$$

where: $\upsilon_{isc,k} = \pi_{kc} + \eta_{iksc}.$

This model – which is equivalent to pooling math and reading and including student fixed effects – is based on several assumptions. Most importantly, it assumes that the effect of teacher numeracy skills on student math performance is the same as the effect of teacher literacy skills on student reading performance. Our data provide support for this assumption as the OLS estimate on teacher numeracy skills is not significantly different from the OLS estimate on teacher literacy skills (p-value = 0.11; see Columns 3 and 6 in Table 2). Another assumption of the within-student model is that any covariate that does not differ across subjects has the same relationship with student reading performance as with math performance. This assumption also does not appear to be critical, because the coefficient on teacher cognitive skills does not change in the student fixed-effects model when we allow for subject-specific impacts of all covariates (results available upon

[26] Within-student across-subject variation has frequently been used in previous research (e.g., Dee (2005, (2007), Clotfelter, Ladd, and Vigdor (2010), Lavy (2015), and Bietenbeck, Piopiunik, and Wiederhold (2017)).

request).[27]

We assume that, conditional on the measured individual, school, and country factors, the differences between math and reading at the country level are random.[28] These differences presumably relate to historical country factors such as the pattern of language development or the historical development of industries and the economy. While difficult to validate, we pursue a variety of approaches designed to uncover significant violations of the key underlying assumption that achievement differences are not simply driven by country patterns in adult competencies.

While the student fixed-effects model accounts for all factors that do not differ between subjects, unobserved differences across countries that are subject-specific (π_{kc} in the error term) remain a potential confounding factor. 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), the student fixed-effects estimates of teacher cognitive skills will still be biased. In Section 5.4, we provide a series of placebo tests and falsification checks, all of which suggest that our teacher-skill estimates do not simply reflect omitted subject-specific factors and that they are not driven by overall population differences in skills. We also address the key issue of separating the impact of teacher skills from the impact of general skills of parents and adults in the country (that might reflect institutions, culture, or other factors).[29]

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

[27] 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.

[28] Supporting this assumption, we find only very low correlations (all magnitudes smaller than 0.1) between the teacher numeracy-literacy skill difference and the following country-level factors: teacher wage premium (see Section 6.2), teacher wage level, share of female teachers, and GDP per capita.

[29] If aggregate school systems differ in systematic ways that increase both teacher cognitive skills and student performance (e.g., a more demanding curriculum), both the OLS estimates and, in case of subject-specific differences, the fixed-effects estimates might be biased. While we cannot directly address this issue, it is reassuring that the pattern of results is robust in various country subsamples with more homogenous quality of educational institutions (see Section 5.3 for details).

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 middle graphs control for the country-specific skills of all adults aged 25–65 to net out the skill persistence across generations.[30] The coefficient on teacher numeracy skills is reduced only modestly, while the coefficient on teacher literacy skills even increases. In the two bottom graphs, we control for all covariates of the baseline OLS specifications (see Table 2 below). While this reduces the teacher-skill estimates, they are still strongly positively associated with student performance.

As expected, the skill level of all adults (aged 25–65) is also strongly positively related to student performance (Figure A-1 in the Appendix). However, when controlling for teacher cognitive skills, the estimates for adult skills substantially decrease in size and lose statistical significance. Below, we show that this aggregate pattern is found in the micro estimates, i.e., the relationship between teacher skills and student performance is not just driven by the overall achievement levels in the country.

5.1 Ordinary Least Squares Results

Table 2 presents our estimates of the relationship between teacher cognitive skills and student performance using student-level test-score data.[31] The unconditional relationship between teacher numeracy skills and individual-level student math performance (Column 1) is identical to the country-level estimate in the upper panel in Figure 3. The coefficient on teacher numeracy skills

[30] 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.

[31] Throughout, we cluster standard errors at the country level because teacher skills do not vary within countries. 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), and 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). Results remain robust when employing the wild bootstrap procedure as an alternative to clustering. We do not make any separate correction for the fact that parent skill is a generated regressor and that this prior estimation might affect the estimated standard errors (Pagan (1984)).

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).[32] The estimate in Column 3 implies that a one SD increase in teacher numeracy skills increases student math performance by almost 0.15 SD. Parent numeracy skills are also significantly related to student performance. The coefficient is rather modest in size compared to teacher cognitive skills, but prior research does not provide any way of making direct comparisons. Importantly, this does not imply that parents have limited impacts. First, the models include a larger set of measures of family background. Second, parent skills might suffer from more measurement error than teacher skills because parent skills are based on information from student reporting of family background.

Columns 4–6 report results for reading. Teacher literacy skills are highly statistically significant across specifications, although the point estimate is somewhat smaller than the coefficient on teacher numeracy skills in the specification with all controls (0.09; see Column 6). Notably, when accounting for student characteristics and family influences (Columns 2 and 5), the point estimates of teacher skills decreases considerably more in reading than in math, suggesting that parents are more important for improving their children's reading abilities than their math performance.[33] Nonetheless, the difference in the math and reading coefficients for teacher skills is not statistically significant.

We find some evidence for heterogeneity of the teacher-skill effect across student subgroups (Table EA-7). The impact of teacher skills is somewhat larger for girls than for boys, for low-SES students compared to high-SES students (particularly in reading), and for natives relative to migrants (particularly in math).[34] Parent cognitive skills are considerably more important for high-

[32] Coefficients on the other control variables are reported in Table EA-6. All coefficients have the expected signs. 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.

[33] Allowing the impact of student characteristics to differ across countries (by interacting the student characteristics with country dummies) does not significantly change the coefficient on teacher skills.

[34] Socioeconomic status (SES) is measured by the PISA index of economic, social, and cultural status (ESCS). Because first-generation migrants might have migrated to the PISA test country just 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

SES students, but there are no differences by student gender or migration status.[35]

To gauge the magnitude of our estimates, we use the OLS coefficients to simulate the improved student performance if each country brought its teachers up to the cognitive-skill level of Finnish teachers, the highest skilled in our sample (Table 3). (This simulation exercise assumes that we have identified a causal effect of teacher cognitive skills.) For Japan, this is not a huge change, yet even Japanese students would improve somewhat (0.06 SD in math and 0.02 SD in reading). But for others, the improvements in student performance would be substantial. U.S. students would be expected to improve by roughly 0.33 SD in math; students in Turkey and Chile, being at the bottom of the international league table, would be expected to improve by about 0.54 SD and 0.57 SD, respectively, in math.

How much would the international differences in student achievement be reduced by improving teacher cognitive skills to the Finnish level? For our 31 countries, the country-level SD of mean PISA scores is 29.3 for math and 21.9 for reading. The simulations in Table 3 imply that bringing teachers in each country to the Finnish level would reduce the country dispersion to 22.1 in math and 15.9 in reading – roughly a reduction by one quarter in each domain.

Of course, moving to the level of Finland is likely to be unrealistic in the short run for many countries. For example, Turkey would have to draw its median teacher from the 97th percentile of the college numeracy distribution instead of the 53rd percentile that it now does (see Table A-1 in the Appendix). For numeracy, nine of the 31 countries would have to increase the place from which the median teacher is drawn by more than 30 percentiles of the distribution of college graduates; for literacy, ten countries would need to move up that far.[36] The U.S. would need to get its median

---

only second-generation migrants when estimating teacher-skill effects for migrants, since these students were born in the PISA test country and have spent their school career in the education system of that country.

[35] We also considered possible heterogeneous impacts of teachers at different parts of the distribution, modeled crudely by including the variance in teacher skills within each country. The variance of teacher skills is statistically insignificant in the student achievement model, and the (median) teacher skills estimates are unaffected. However, because of the small teacher samples, we are concerned that we do not have a good description of the distribution of teacher skills in each country. Interpretation is also clouded by the possibility that a larger variance of teacher skills implies greater measurement error at the individual student level.

[36] In numeracy, greater than a 30 percentile change to meet Finland teachers is required in Chile, Estonia, Israel, Italy, Korea, Poland, Russian Federation, Spain, and Turkey. In literacy, greater than a 30 percentile change to meet

math (reading) teacher from the 74th (71st) percentile instead of the current 47th (51st) percentile.

To understand the magnitude of the estimated impact of teacher skills, it is important to note that the teacher-skill estimates do not capture the effect of just one school year but rather reflect the cumulative effect of teacher cognitive skills on student performance over the first ten school years. Thus, these projections are long-run impacts that presume that the quality of students' teachers across the first ten grades would improve to the level of Finland.

5.2 Student Fixed-Effects Results

While the previous section has shown that teacher cognitive skills are significantly related to student performance in both math and reading, the possibility of country-specific omitted variables remains. Therefore, we now exploit only within-country variation to identify the effect of teacher cognitive skills on student performance, eliminating any non-subject-specific bias.

Again, the overall story is easy to see in a simple diagram. Aggregating student performance to the country level, Figure 4 shows that *differences* in teacher cognitive skills between numeracy and literacy are systematically related to *differences* in student performance between the same two subjects.

Table 4 presents the results of the student fixed-effects specifications using student-level test score data. Here, all control variables that differ across subjects are included in first differences, while all factors common to the two subjects drop out. Across specifications, the student fixed-effects estimates for teacher cognitive skills of 0.11 remain sizeable[37] and close to the OLS coefficients on teacher numeracy and literacy skills.[38] While neither parent cognitive skills nor teacher shortages (which differ by subject) are significantly related to student performance in the

Finland teachers is required in Austria, Chile, Denmark, Israel, Italy, Lithuania, Russian Federation, Slovak Republic, Slovenia, and Turkey.

[37] Replacing median teacher cognitive skills with alternative measures at other parts of the skill distribution, such as the 25th percentile and the 75th percentile, or with the fraction of teachers whose numeracy skills exceed their literacy skills, also yields a positive relationship with student performance.

[38] Using the Stata command *suest*, we find that the coefficient on teacher cognitive skills from the student fixed-effects model is not statistically significantly different from the OLS coefficients (p-value = 0.40 in math and p-value = 0.78 in reading). The comparisons refer to the full-control specifications (i.e., Column 3 of Table 4 and Columns 3 and 6 of Table 2).

student fixed-effects models, the effect of instructional time on student performance is significant and similar to the effect size in Lavy (2015).[39]

These results, which are not subject to the same potential biases as the previous OLS estimates, strongly support the role of teacher skills. Moreover, the consistency across estimation approaches in magnitude and significance is notable.

5.3 Robustness Checks

In this section, we show that our results are robust to controlling for country-level skill differences in various ways, to including additional country-level controls, and to using different country subsamples.

Overall Country-Level Skill Differences

One concern in the previous models is that the estimated impacts of teacher cognitive skills may reflect subject-specific country differences. In Table 5, we therefore additionally control for the country's cognitive-skill level of parents and of all adults to account for countries' potential subject preference or other subject-related differences (for comparison, Columns 1 and 4 present the baseline models).[40] Adding these cognitive-skill levels does not substantially weaken the impact of teachers' cognitive skills on student performance (see Columns 2 and 3 for math and Columns 5 and 6 reading)..

Teacher cognitive skills reflect both a country's overall cognitive-skill level and where teachers are drawn from the country's skill distribution. In alternative estimation, holding constant the skill level of adults, students perform better in countries where teachers are drawn from further up the cognitive skill distribution (Table EA-8). This provides additional evidence that the estimates on teacher cognitive skills are not driven merely by international differences in overall cognitive skills. It matters from where in the pool of potential teachers countries draw their teachers.

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

[40] 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).

Controlling for Additional Country-Level Factors

Teachers' subject-specific cognitive skills might be correlated with their subject-specific pedagogical skills, implying that the estimated coefficient on teacher skills might partially pick up the impact of pedagogical skills. Using information from the PISA students about their teachers' activities in language and math classes, we construct indicators of subject-specific instructional activities as proxies for teachers' pedagogical skills (see Electronic Appendix A). Controlling for the instructional practices in math and language classes does not affect the teacher-skill estimates, which supports the independent impact of teacher cognitive skills on student performance.[41]

Moreover, the estimates on teacher cognitive skills remain significant when controlling for GDP per capita (as a measure of a country's state of development) and central exit exams (reflecting a student performance-enhancing institution).

Country Subsamples

The teacher-skill effect holds across different subsamples (Table EA-10). To address divergent cultures (especially differing educational attitudes), we include continental fixed effects (and alternatively restrict the analysis to just European countries). Furthermore, we exclude ex-communist countries (the Czech Republic, Estonia, Lithuania, Poland, Russia, the Slovak Republic, and Slovenia) and Turkey, where occupational choices were less driven historically by market incentives and often depended on political attitudes. Finally, we use only countries with larger teacher PIAAC samples, where measurement error in country-level teacher cognitive skills is likely smaller.[42] The teacher-skill coefficient is similar in all these alternative specifications, lying within the 95% confidence interval of the baseline estimate.

Moreover, using the baseline OLS specification with all control variables, excluding each

[41] See Table EA-9 for the results when controlling for instructional practices and for the other country factors described below.

[42] Measurement error generally has a more severe impact in fixed effects models. When applying the same sample restriction in the student fixed-effects model, the coefficient on teacher cognitive skills increases somewhat, suggesting that using only countries with larger teacher PIAAC samples likely reduces measurement error in the subject differences of teacher cognitive skills.

country individually from the sample yields teacher-skill estimates that are always very close to the baseline coefficients, indicating that results are not driven by any individual country (results not shown).

5.4 Placebo Tests and Falsification Checks

Placebo Tests

While the teacher-skill estimates remain statistically significant and vary little across specifications and estimation methods, it is clearly difficult to guard against all possible omitted country-level factors that could bias the results. We specifically remain concerned that student performance is positively related to adult skills simply because countries with high adult skills also have teachers with high skills.

Therefore, we perform various placebo tests, all of which indicate that our estimates reflect the impact of teacher cognitive skills and not just those of the society in general.

In the first placebo test, we replace teacher cognitive skills with the cognitive-skill level of workers in occupations other than teaching. For this analysis, we use all occupations with at least 100 observations on average across PIAAC countries, resulting in 14 occupations that cover the full range of a country's occupational distribution (e.g., managers, scientists and engineers, health professionals, business professionals, clerks, sales workers, service workers). To address the concern that skills of workers in these other occupations simply capture the overall skill level of a country, we control for the cognitive-skill level of adults. (Note that we do not control for adult skills in the student fixed-effects model because general skill differences across countries – as well as all other subject-invariant differences – are already accounted for.)

Figure 5 depicts the estimated coefficients on the cognitive skill-levels of workers in each of the other occupations. For comparison, the vertical dashed lines indicate the estimated impacts of teacher cognitive skills. The left and middle graphs show the OLS results. Student performance in math (reading) is significantly positively related to the skills of workers in only one occupation (two occupations). Moreover, out of the 28 OLS coefficients on numeracy and literacy skills, only one

coefficient is larger than the respective estimate on teacher skills.

The right graph of Figure 5 presents the student fixed-effects estimates on the cognitive-skill level in other occupations. While three coefficients are statistically significant, the skill level in only one occupation – health workers – is more strongly related to student performance than are teacher cognitive skills. Importantly, from the OLS models in the other two panels, the cognitive skills of health workers are actually negatively related to student math and student reading performance. Overall, there is no occupation other than teaching whose skill level is systematically related to student performance across the OLS and student fixed-effects models.[43]

In a second placebo test, we replace teacher skills by the skill level of a randomly chosen sample of adults matched by age, gender, and educational distribution to the teacher sample in each country and having the same sample size as the country-specific teacher sample. In each country, we draw 100 samples of matched "teacher twins", thus comparing 100 coefficients with the estimate for teacher cognitive skills. Figure 6 shows histograms of estimated student impacts (with the dashed line again showing the estimated coefficient for the actual sample of teachers). Only 9% of the numeracy-skill estimates, 1% of the literacy-skill estimates, and 2% of the student fixed-effects estimates exceed the respective estimate on teacher cognitive skills. Importantly, *none* of the 100 samples of teacher twins produces consistently larger skill coefficients than teachers in all three models. Thus, both placebo tests provide strong evidence that the estimates on teacher cognitive skills are not systematically biased by unobserved country-level skill differences.

Falsification Checks

We also investigate cross-subject effects, i.e., the effect of teachers' numeracy skills on student reading performance and the effect of teacher literacy skills on student math performance. If it is subject-matter skills that are important, as we have assumed, teacher skills in one subject would be

[43] Additionally, we estimate specifications that include teacher cognitive skills *and* simultaneously the cognitive skill levels in all 14 other occupations. In these horse-race-type models, teaching is the only occupation that is significantly positively related to student performance in both OLS and student fixed effects (results available on request).

only weakly related – if at all – to student performance in the other subject. (A positive relationship may arise due to cross-subject spillover effects, that is, higher literacy skills make teachers also better at teaching math and vice versa.)

This estimation supports the importance of subject-specific skills (Table A-2). Teacher numeracy skills have a substantially larger association with student math performance than with student reading performance. Similarly, teacher literacy skills are more important for student reading performance than for student math performance once we account for general country-level differences in adult skills. The most convincing evidence comes from simultaneously including teacher numeracy and literacy skills. Here, teacher skills in either subject only affect student performance in the same subject; teacher skills in the other subject are always close to zero and statistically insignificant.

In a final falsification check, we use the measure of teachers' ICT skills, the third skill domain tested in PIAAC, instead of teacher skills in numeracy or literacy (Table A-3). Conditioning on the overall skill level in the country, we find that teacher ICT skills are not significantly related to student performance in either math or reading. (However, note that standard errors are larger for teacher ICT skills than for teacher numeracy or literacy skills; see Table 2.) Thus, it is the subject-specific skills of teachers that affect student performance (in the same subject), and not simply general differences in teacher quality across countries.

6. Determinants of Teacher Cognitive Skills

International differences in teacher cognitive skills reflect both where teachers are drawn from each country's skill distribution and the overall skill level of each country's population – and policies to improve the skills of teachers could conceptually focus on either of these dimensions. Increasing the overall skill level of a country's population would of course be both desirable and self-reinforcing through improving the pool of potential teachers. Nonetheless, potential overall population improvement policies, while widely discussed elsewhere, are 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 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 and ultimately student outcomes.

6.1 Alternative Professional Opportunities for Women

We first investigate whether teacher skills are affected by competition from other occupations that demand high skills. To do so, we explore changes in alternative job opportunities for women over time.[44] 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. The underlying idea is that historically women have been segregated into a constrained set of occupations, one of which is teaching. While the causes of segregation are not well-tested, the empirical reality is that women previously were much more concentrated in teaching than they generally are today (see Bergmann (1974), Temin (2002), Lakdawalla (2006)).

Our analysis, however, differs from Bacolod (2007) in two key ways. First, we explicitly consider the human capital intensity of alternative occupations, rather than relying on relative average wages in teaching and elsewhere. Second, we observe multiple countries, which dramatically expands the range of observations and permits accounting for any general (i.e., non-country-specific) time trends that affect both female labor-market participation and teacher skills. For example, the teaching profession might have become less attractive relative to other high-skill occupations over time, explaining both an increasing share of females in other high-skill occupations and a decline in average teacher skills.

To measure women's access to high-skill occupations, we compute the proportion of female teachers relative to the number of females in all high-skill occupations. Using the PIAAC micro

[44] We focus on labor market opportunities for women because they constitute the majority in the teacher workforce. Across the 23 countries used in the analysis below, more than two-thirds (69 percent) of teachers are female, ranging from 59 percent in Japan to 79 percent in Austria. Note that the gender composition of the teacher sample does not affect the numeracy-literacy skill gap (r = -0.06 in the 23 country sample and r = -0.1 in the full sample).

data, we determine "high-skill" occupations empirically in each country. First, we calculate the average years of schooling of employees working in each two-digit occupation at the time of the PIAAC assessment.[45] Second, ranking occupations in each country by average schooling level in descending order, we define all occupations as "high-skill" until the top-ranked occupations comprise 25 percent of all working males in the country.[46] To obtain groupings with sufficient numbers of teachers, we combine 15 adjacent age cohorts. Since the PIAAC data cover 45 birth years (excluding very young adults who mainly have not completed college education), this yields three aggregated age cohorts in each of the 23 countries used in this analysis.[47]

To test whether a higher concentration of females in teaching is associated with higher cognitive skills of teachers, we estimate the following model:[48]

$$T_{kca} = \alpha + \beta Share_{ca} + \mu_a + \mu_c + \gamma College_{kca} + \epsilon_{kca}, \quad (6)$$

where T_{kca} denotes median teacher skills in subject *k* in country *c* in age cohort *a* and $Share_{ca}$ is the share of female teachers relative to all women working in high-skill occupations in country *c* in aggregate age cohort *a*. We always include cohort fixed effects (μ_a) to control for general time trends in women's labor-market opportunities and for skill depreciation over the lifecycle. Moreover, country fixed effects (μ_c) account for cross-country differences in women's labor-market participation and in average skill levels that are constant across birth cohorts. In some specifications, we additionally control for the cognitive skills of nonteacher college graduates in the

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

[46] Note that teaching is a high-skill 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 skill leads to qualitatively similar results. The 25-percent rule ensures that a similar share of workers is employed in high-skill occupations in each country; other variants of defining high-skill occupations led to more uneven shares of males working in high-skill occupations across countries.

[47] Since our analysis uses only pseudo cohorts based on the cross-sectional PIAAC data, the validity of our results depends on the assumption that women do not change the type of their occupation (high-skill 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.

[48] We exclude all ex-communist countries and Turkey since occupational choices in these countries were less driven by market incentives but rather depended on political attitudes.

respective country-cohort cell to account for country-specific skill depreciation ($College_{kca}$).[49]

For both numeracy and literacy, we find that more females working in teaching relative to other high-skill occupations is significantly related to higher teacher cognitive skills (Table 6).[50] The estimates are also economically meaningful. While somewhat smaller for literacy, an increase by 10 percentage points in the share of female teachers over all females working in high-skill occupations is associated with a 0.36 SD increase in the numeracy skills of teachers. The average share of female teachers over all females working in high-skill occupations across all three cohorts varies between 17 percent in Chile (18 percent in the U.S.) and 38 percent in Singapore. Thus, if females in the U.S. had similar employment choices and opportunities as in Singapore, average teacher numeracy skills in the U.S. would be about 0.72 SD higher, bringing U.S. teachers to just above the international average in teacher numeracy skills. Across all 23 countries in the sample, the share of female teachers over all females working in high-skill occupations decreases from 29 percent in the oldest age cohort (born in years 1946–1960) to 22 percent in the youngest cohort (born 1976–1990), reflecting an international improvement of alternative job opportunities for women over time. This is associated with an international decline in teacher numeracy skills by 0.25 SD.

6.2 Teacher Pay

An obvious consideration in looking at the pattern of teacher skills is the pay of teachers. 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.[51] If this link

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

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

[51] 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 effects is, however, not very strong; 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 just at effort; see Woessmann (2011). For developing countries, the evidence on effort is stronger (see Muralidharan and Sundararaman

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.[52]

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-type earnings equation with log earnings (ln *y*) regressed on gender (*G*), potential work experience (*E*), achievement in numeracy and literacy (*A*), and a binary teacher indicator (*T*):[53]

$$\ln y = a_0 + a_1 G + a_2 E + a_3 E^2 + \boldsymbol{A}\boldsymbol{a}_4 + \delta T + \varepsilon. \quad (7)$$

The coefficient δ is the earnings premium (in percentage terms) for teachers given their characteristics. We estimate a separate premium for each country, and we find a wide dispersion. Figure 7 shows the estimated teacher premiums across countries, ranging from +45 percent in Ireland to -22 percent in the United States and Sweden.[54] (Table EA-11 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). Our estimates indicate that teachers are paid some 20 percent less than *comparable* college graduates elsewhere in the U.S. economy, that is, after adjusting for observable characteristics.

We now put teacher pay and teacher skills together by regressing cognitive skills of teachers (in

(2011)), but this might not generalize to the developed countries we analyze. Among other things, the very high rates of teacher absenteeism in many developing countries indicates more room for improvement on the effort margin.

[52] 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.

[53] This approach follows Hanushek et al. (2015, (2017) 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.

[54] 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, the 90/10 wage ratio is 4.5 compared to 2 in Sweden.

subject *k* in country *c*) on the country-specific teacher wage premium ($\hat{\delta}_c$):

$$\bar{T}_{kc} = \alpha + \beta\hat{\delta}_c + \gamma College_{kc} + \epsilon_{kc}. \quad (8)$$

Estimates are conditioned on the cognitive skills of nonteacher college graduates to account for 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.[55]

The results, shown in Table 7 (and graphically in Figure 8) 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 tend to draw their teachers from higher parts of the college 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.[56] The coefficient on college graduates' skills is close to 1, again suggesting the powerful influence of a country's overall skill level.

It is also possible to estimate a "reduced-form" effect of teacher wage premium on student performance. The lower panel of Figure 8 shows that the pay choices of countries carry through to student achievement.

The interpretation of these results is, however, important for policy. These estimates are reduced-form estimates that reflect the labor market equilibrium. As such, they do not 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.[57] Moreover,

[55] 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 other earnings characteristics between teachers and nonteachers. Hoxby and Leigh (2004) suggest that compression in salaries, as caused by unionization, also might contribute to differences in teacher cognitive skills, but we have no way to incorporate such differences into our estimation.

[56] 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.

[57] 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 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

the estimated relationship relates to the long run after many cohorts of teachers have been recruited. Thus, while making it clear that a more skilled teaching force is generally found in countries with higher relative salaries, the evidence says nothing about either how salaries should be structured or the responsiveness of teachers to higher salary offers.

7. Conclusions

In this paper, we investigate the role of teacher skills in explaining international differences in student achievement. Within-country evidence has highlighted the importance of teacher quality for student achievement with the most convincing evidence coming from value-added analysis. Such analysis provides information about the relative learning gains across a set of teachers, but it does not indicate what might be possible if there were a different pool of potential teachers from which the teacher corps could be drawn. Moreover, it has previously not been possible to describe reliably any aspects of teachers that could be used to index quality differences across countries.

Based on suggestive prior evidence on the role of cognitive skills of teachers, we systematically address how cross-country differences in teacher skills enter into educational production. We use newly available data from the Programme for the International Assessment of Adult Competencies (PIAAC) to provide the first description of the skills of teachers in numeracy and literacy. For our sample of 31 developed economies, teacher cognitive skills differ substantially across countries, reflecting both country-wide differences in cognitive skills and policy choices about where teachers are drawn from the country's skill distribution. 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. These estimates account for a rich set of controls for student, school, and country background factors, including coarse measures of the cognitive skills of the parents of PISA students and a variety of institutional features of the schools in each country. In addition to OLS models, we estimate the impact of teacher cognitive skills using student fixed-

induced to stay in teaching are the more effective teachers. The same holds for changing the cognitive skills of the teaching force. See Hanushek and Rivkin (2004).

effects models, which exploit between-subject variation and account for constant individual factors (e.g., ability and parental influences) along with non-subject-specific country-level factors.

With both approaches, we consistently find that differences in teacher cognitive skills across countries are strongly associated with international differences in student performance. In terms of magnitude, a one SD increase in teacher cognitive skills is associated with an increase in student performance of 0.10-0.15 SD. 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.

Alternative specifications that control for the cognitive skills of all adults in a country 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. Several placebo tests and falsification checks support a conclusion that the estimates of the relationship of student outcomes and teacher cognitive skills are unlikely to be driven by overall population differences in skills or by omitted country variables.

The magnitude of the estimated relationship is important. Our results suggest that the dispersion in average PISA scores across our 31 country sample would be reduced by roughly one-quarter if each country brought its average teacher skills up to the average in Finland, the country with the highest measured skills of teachers. Of course, reaching Finland has very different labor market implications across countries because of underlying differences in the skills found across countries in the pool of college graduate. For example, Chile and Turkey would need to draw their median teacher at the 97th percentile of the distribution of numeracy scores for college graduates in order to reach the level of Finnish teachers.

We also investigate possible determinants of teacher cognitive skills. We find that cross-country differences in women's access to high-skill occupations and in wage premiums paid to teachers (given their gender, work experience, and cognitive skills) are directly related to teacher cognitive skills in a country. The estimated wage differentials for teachers are directly correlated with student outcomes across our sample countries.

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.

Bietenbeck, Jan, Marc Piopiunik, and Simon Wiederhold. Forthcoming. "Africa's skill tragedy: Does teachers' lack of knowledge lead to low student performance?" *Journal of Human Resources*.

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. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41, no. 4 (Fall): 778-820.

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

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. 2010. "If you pay peanuts do you get monkeys? A cross country analysis of teacher pay and pupil performance." Mimeo. London: Royal Holloway College, University of London.

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.

Ferguson, Ronald F., and Helen F. Ladd. 1996. "How and why money matters: An analysis of Alabama schools." In *Holding schools accountable: Performance-based reform in education*, edited by Helen F. Ladd. Washington, DC: Brookings: 265-298.

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.

Goldhaber, Dan. 2007. "Everyone's Doing It, But What Does Teacher Testing Tell Us About Teacher Effectiveness?" *Journal of Human Resources* XLII, no. 4 (Fall): 765-794.

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. 1971. "Teacher characteristics and gains in student achievement: Estimation using micro data." *American Economic Review* 60, no. 2 (May): 280-288.

Hanushek, Eric A. 1972. *Education and race: an analysis of the educational production process*. Cambridge, MA: Health-Lexington.

Hanushek, Eric A. 1992. "The trade-off between child quantity and quality." *Journal of Political Economy* 100, no. 1 (February): 84-117.

Hanushek, Eric A. 1995. "Interpreting recent research on schooling in developing countries." *World Bank Research Observer* 10, no. 2 (August): 227-246.

Hanushek, Eric A. 1997. "Assessing the effects of school resources on student performance: An update." *Educational Evaluation and Policy Analysis* 19, no. 2 (Summer): 141-164.

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. 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: 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. 2004. "How to improve the supply of high quality teachers." In *Brookings Papers on Education Policy 2004*, edited by Diane Ravitch. Washington, DC: Brookings Institution Press: 7-25.

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. 2015. "Returns to skills around the world: Evidence from PIAAC." *European Economic Review* 73: 103-130.
Hanushek, Eric A., Guido Schwerdt, Simon Wiederhold, and Ludger Woessmann. 2017. "Coping with change: International differences in the returns to skills." *Economic Letters* 153(April): 15-19.
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.
Harden, Jeffrey J. 2011. "A Bootstrap Method for Conducting Statistical Inference with Clustered Data." *State Politics & Policy Quarterly* 11, no. 2 (June): 223-246.
Heckman, James J., and Tim Kautz. 2012. "Hard evidence on soft skills." *Labour Economics* 19, no. 4: 451-464.
Hoxby, Caroline M., and Andrew Leigh. 2004. "Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States." *American Economic Review* 94, no. 2 (May): 236-240.
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).
Jackson, C. Kirabo. 2012a. "Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers." *Review of Economics and Statistics* 95, no. 4 (October): 1096-1116.
Jackson, C. Kirabo. 2012b. "School competition and teacher labor markets: Evidence from charter school entry in North Carolina." *Journal of Public Economics* 96, no. 5–6 (June): 431-448.
Jackson, C. Kirabo. forthcoming. "What do test scores miss? The importance of teacher effects on non-test-score outcomes." *Journal of Political Economy*.
Jackson, C. Kirabo, and Elias Bruegmann. 2009. "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." *American Economic Journal: Applied Economics* 1, no. 4 (October): 85-108.
Jackson, C. Kirabo, Jonah E. Rockoff, and Douglas O. Staiger. 2014. "Teacher effects and teacher related policies." *Annual Review of Economics* 6: 801-825.
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.
Lakdawalla, Darius. 2006. "The Economics of Teacher Quality." *The Journal of Law & Economics* 49, no. 1 (April): 285-329.
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.
Lee, Jong-Wha, and Robert J. Barro. 2001. "Schooling quality in a cross-section of countries." *Economica* 68, no. 272: 465-488.
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. 2010a. *PISA 2009 Results: What Makes a School Successful? – Resources, Policies and Practices*. Vol. IV. Paris: Organisation for Economic Cooperation and Development.

OECD. 2010b. *PISA 2009 results: What students know and can do – Student performance in mathematics, reading and science (Volume I)*. Paris: Organisation for Economic Co-operation and Development.

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

Pagan, Adrian. 1984. "Econometric Issues in the Analysis of Regressions with Generated Regressors." *International Economic Review* 25, no. 1 (February): 221-247.

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.

Rockoff, Jonah E., Brian A. Jacob, Thomas J. Kane, and Douglas O. Staiger. 2011. "Can you recognize an effective teacher when you recruit one?" *Education Finance and Policy* 6, no. 1 (Winter): 43-74.

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.

Temin, Peter. 2002. "Teacher Quality and the Future of America." *Eastern Economic Journal* 28, no. 3 (Summer): 285-300.

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., Matthew A. Kraft, Amy S. Finn, Rebecca E. Martin, Angela L. Duckworth, Christopher F. O. Gabrieli, and John D. E. Gabrieli. 2016. "Promise and Paradox: Measuring Students' Non-Cognitive Skills and the Impact of Schooling." *Educational Evaluation and Policy Analysis* 38, no. 1 (March): 148-170.

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. 2005. "Educational production in Europe." *Economic Policy* 20, no. 43: 446-504.

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

Electronic Appendix A. Data Appendix

This electronic appendix provides additional information on the datasets and the construction of variables used in the empirical analysis. As in Section 3, we begin with a discussion of teacher cognitive skills, followed by a description of parent cognitive skills, student performance data, and further control variables.

Teacher Cognitive Skills

The target population of the Programme for the International Assessment of Adult Competencies (PIAAC) was the non-institutionalized population aged 16-65 years, and samples included at least 5,000 participants in each country.

We use 31 out of the 33 countries that participated in PIAAC 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, conducted between August 2011 and March 2012) as well as Chile, Greece, Indonesia, Israel, Lithuania, New Zealand, Singapore, Slovenia, and Turkey (Round 2, conducted between April 2014 and March 2015). Cyprus, while participating in PIAAC, did not participate in PISA. In Indonesia, the PIAAC survey was administered only 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.

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 survey mode was to answer questions on a computer, but respondents without computer

experience could opt for a pencil-and-paper interview. On average across countries, 70 percent of the 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).

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 (or ICT skills). 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.[58]

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" in the main analysis.[59]

In the PIAAC 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

[58] 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.

[59] Five countries surveyed in PIAAC (Cyprus, France, Indonesia, Italy, and Spain) did not administer tests in this optional skill domain.

Standard Classification of Occupations) codes for all but two countries, which allows us to identify teachers in fine categories. For Australia and Finland, we only have two-digit occupational codes and are therefore 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 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.

Because PIAAC is not a teacher survey, we benchmark the PIAAC teacher samples against large administrative datasets in which detailed occupational information allows identifying teachers. Using the 2011 U.S. American Community Survey (ACS) which includes 55,000 teachers and the 2011 German Micro Census which includes about 6,400 teachers, we find similar demographic characteristics compared to the respective national PIAAC teacher samples. In the U.S. teacher samples, mean age is 41.1 years in PIAAC (44.3 years in the ACS), 67.5% (75.5%) of teachers are female, and 89.9% (88.6%) of teachers have a college degree. In the German teacher samples, mean age is 47.2 years in PIAAC (45.6 years in the Micro Census), 65.8% (74.4%) of teachers are female, and 85.2% (82.2%) of teachers have a college degree.

Parent Cognitive Skills

When estimating the numeracy and literacy skills of the PISA parents from the PIAAC micro data, 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 EA-1).

Student Performance and Further Control Variables

The Programme for International Student Assessment (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 students 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 focus subject 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, a potentially important confounding factor for teacher cognitive skills. Alternatively, the subject-specific pedagogical skills might also mediate the impact of teacher cognitive skills if teachers with high skills also use better instructional practices in class (see Section 5.3).

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. We use estimated parent cognitive skills, number of books at home, parents' highest educational degree, and parental occupation to control for family background (see Table A-2 for summary statistics of student performance and student characteristics and Table A-3 for family

background controls).

Based on student information, we can construct measures of weekly instructional time for both language and math classes. Following Lavy (2015), we aggregate this information across students to the school level. Following Hanushek, Link, and Woessmann (2013), we also control for various school-level determinants of student performance. For example, 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).

To construct indicators of subject-specific instructional activities, we use information from the PISA students about their teachers' activities in language and math classes. We follow the OECD (2010a) approach of measuring specific instructional practices through survey responses of students, while we aggregate these instructional practices to the school level. For *reading*, we use the following items (each measured on a 4-point scale ranging from "never or hardly ever" to "in all lessons"): 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 how they have solved a problem; and presenting problems that require students to apply what they have learnt to new contexts. Since instructional practices are asked only for the subject that was the focus in the respective PISA cycle (reading in 2009 and math in 2012), we impute the subject-specific instructional-practice indicator for the PISA cycle when a subject was not the focus. We impute the 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.

Electronic Appendix B. 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 cohorts 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.[60]

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.[61] 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

[60] 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.

[61] 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.

(literacy).

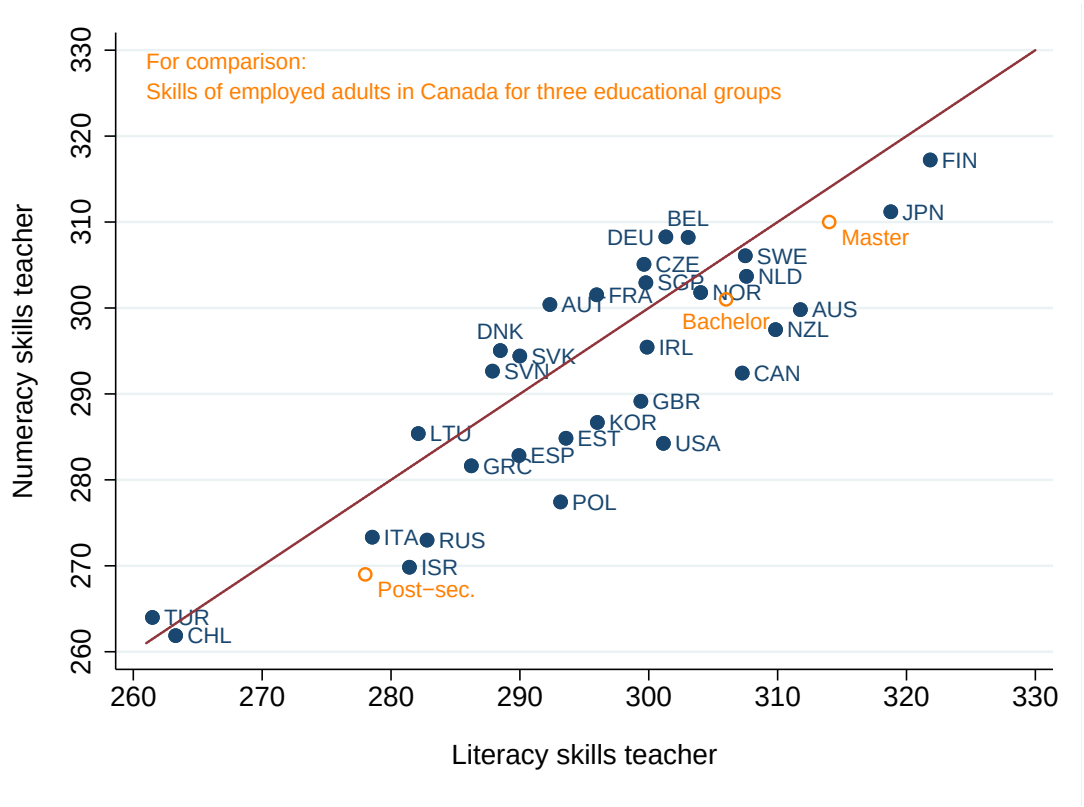
We also compare teacher cognitive skills from PIAAC with those from Germany's adult cohort of the National Educational Panel Study (NEPS).[62] This dataset is a nationally representative dataset of 9,352 adults born between 1944 and 1986. NEPS has several advantages for our purpose. First, similar to 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 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.

[62] 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).

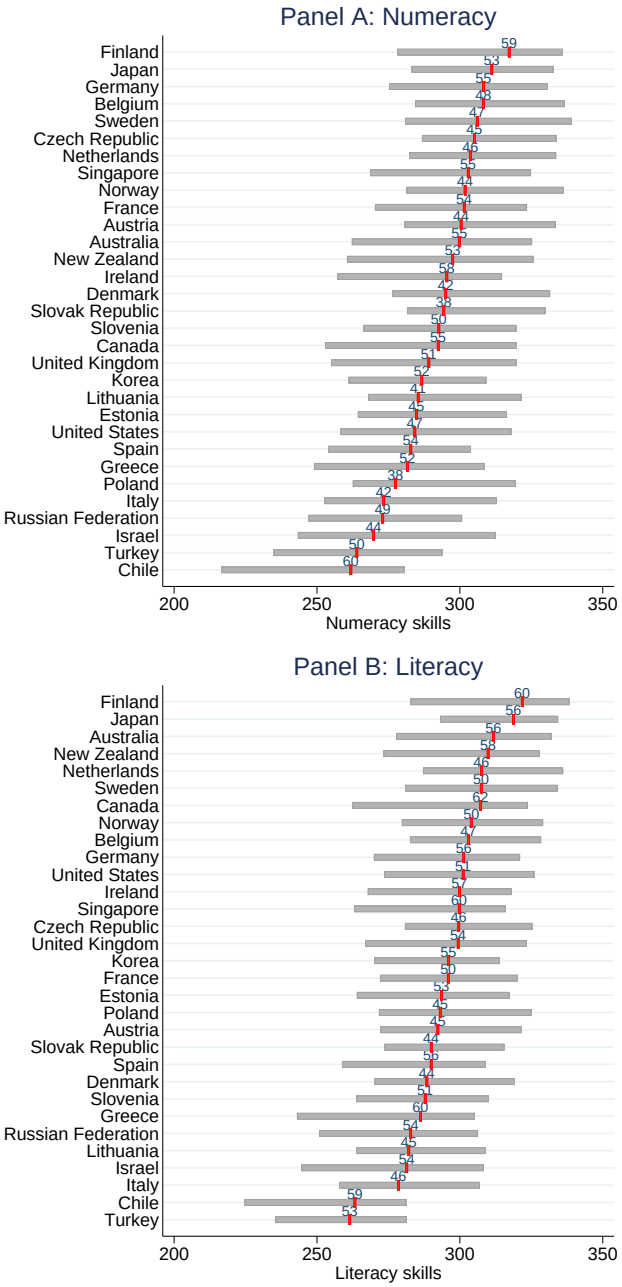
Figures and Tables

Figure 1: Teacher Cognitive Skills Compared to Canadian Workers with Varying Education Levels



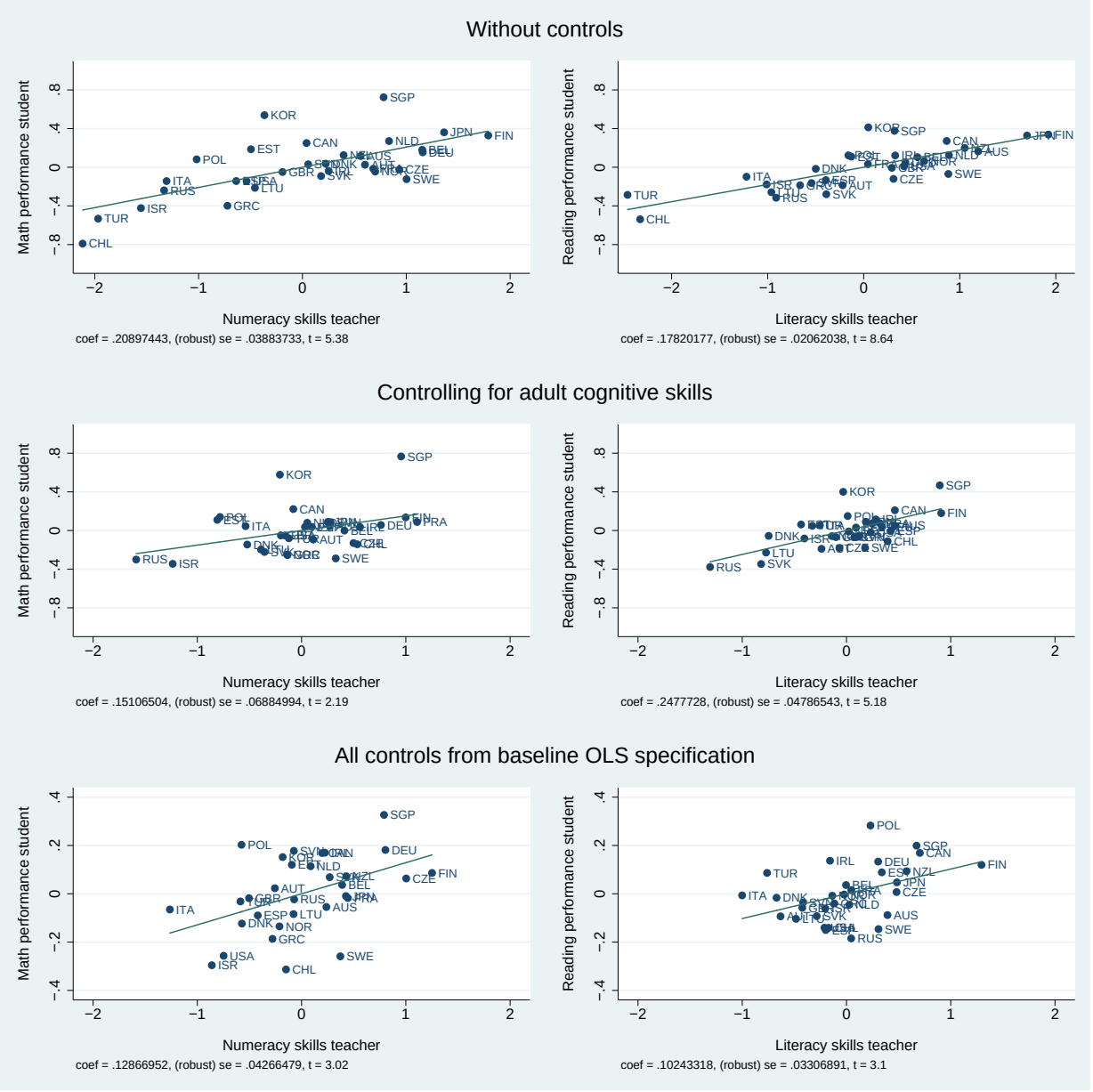
Note: The blue dots indicate country-specific teacher skills in numeracy and literacy (see text for construction of teacher cognitive skills). The orange circles indicate the median cognitive skills for three educational groups of employed adults aged 25–65 years in Canada (the largest national sample in PIAAC). *Post-sec.* includes individuals with vocational education (post-secondary, non-tertiary) as highest degree (2,434 observations); *Bachelor* includes individuals with bachelor degree (3,671 observations); *Master* includes individuals with a master or doctoral degree (1,052 observations). *Data sources*: PIAAC 2011/12 and 2014/15.

Figure 2: Position of Teacher Cognitive Skills in the Skill Distribution of College Graduates



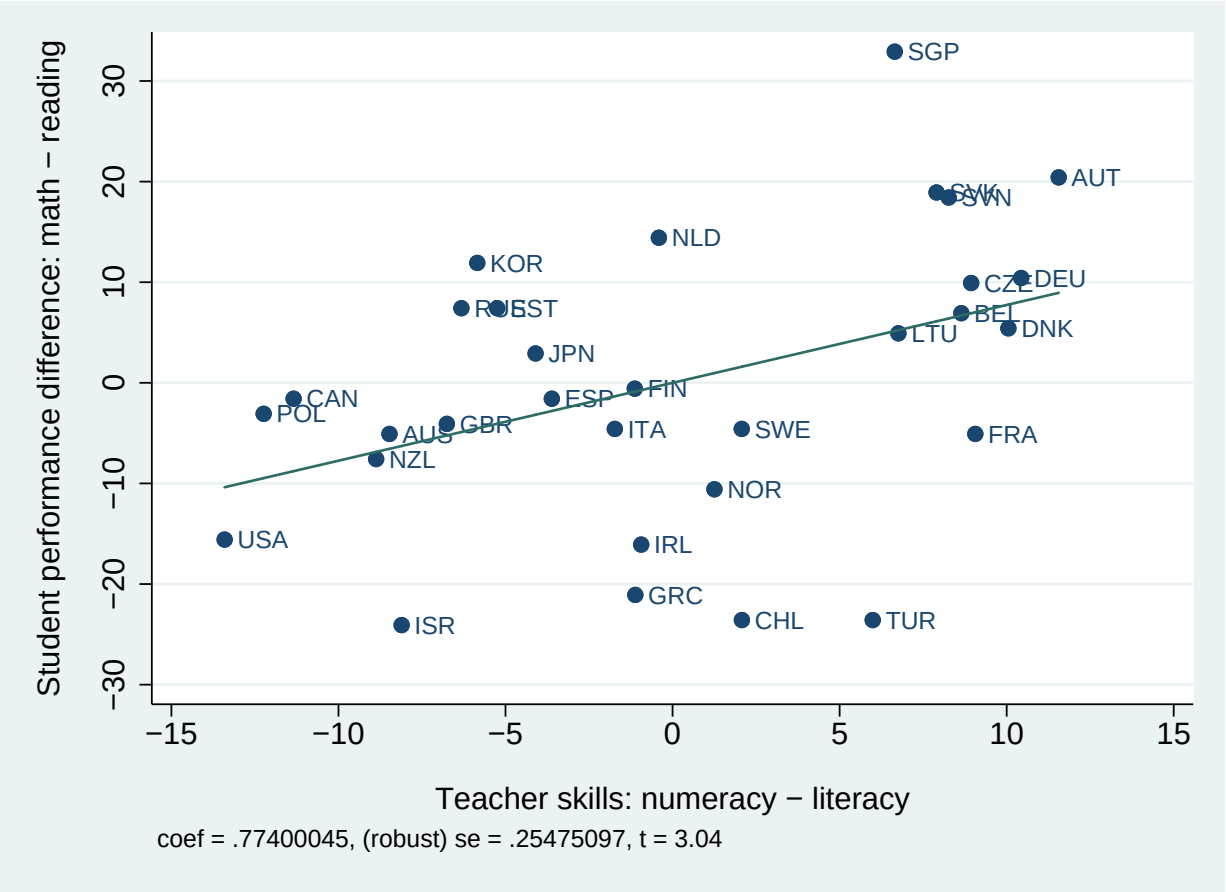
Note: Modified figure from Schleicher (2013). The vertical red bars indicate the median cognitive skills of teachers in a country. Horizontal bars show the interval of cognitive skill levels of all college graduates (including teachers) between the 25th and 75th percentile. Numbers on top of the vertical bars indicate the position of teacher cognitive skills in the cognitive skill distribution of college graduates. Countries are ranked by the median teacher skills in numeracy and literacy, respectively. *Data sources*: PIAAC 2011/12 and 2014/15.

Figure 3: Student Performance and Teacher Cognitive Skills



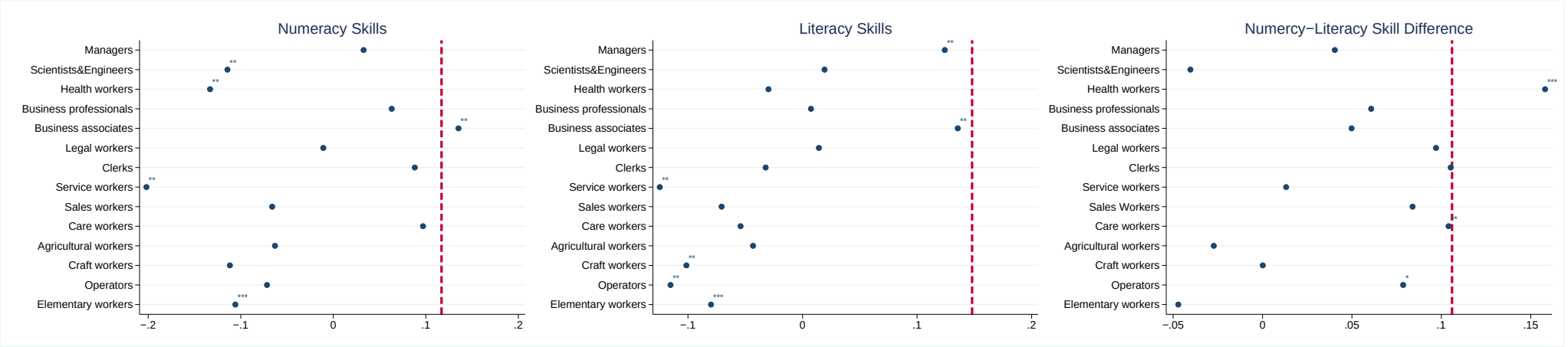
Note: The two graphs in the top panel do not include any controls. The two graphs in the middle panel are added-variable plots that control for country-specific average skills in numeracy and literacy, respectively, of all adults aged 25–65. The two plots in the bottom panel are added-variable plots that control for all variables included in the baseline OLS specification in Columns 3 and 6 of Table 2. *Data sources*: OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Figure 4: Student Performance Difference and Teacher Cognitive Skills Difference



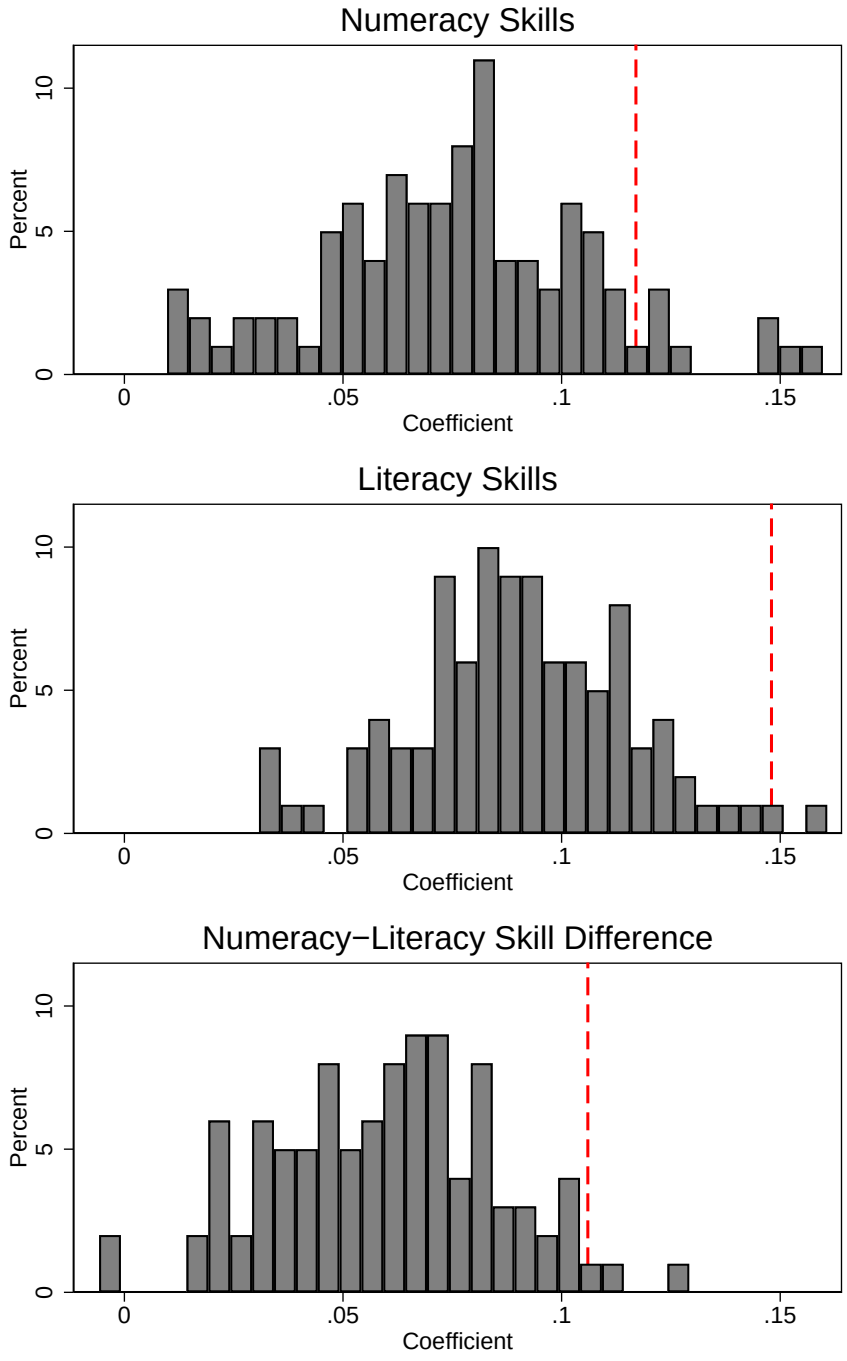
Note: The graph plots the student performance difference between math and reading (at the country level) against the difference in teacher cognitive skills between numeracy and literacy (at the country level). *Data sources*: PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Figure 5: Placebo Tests Using Cognitive Skills in Other Occupations (OLS and Student Fixed Effects)



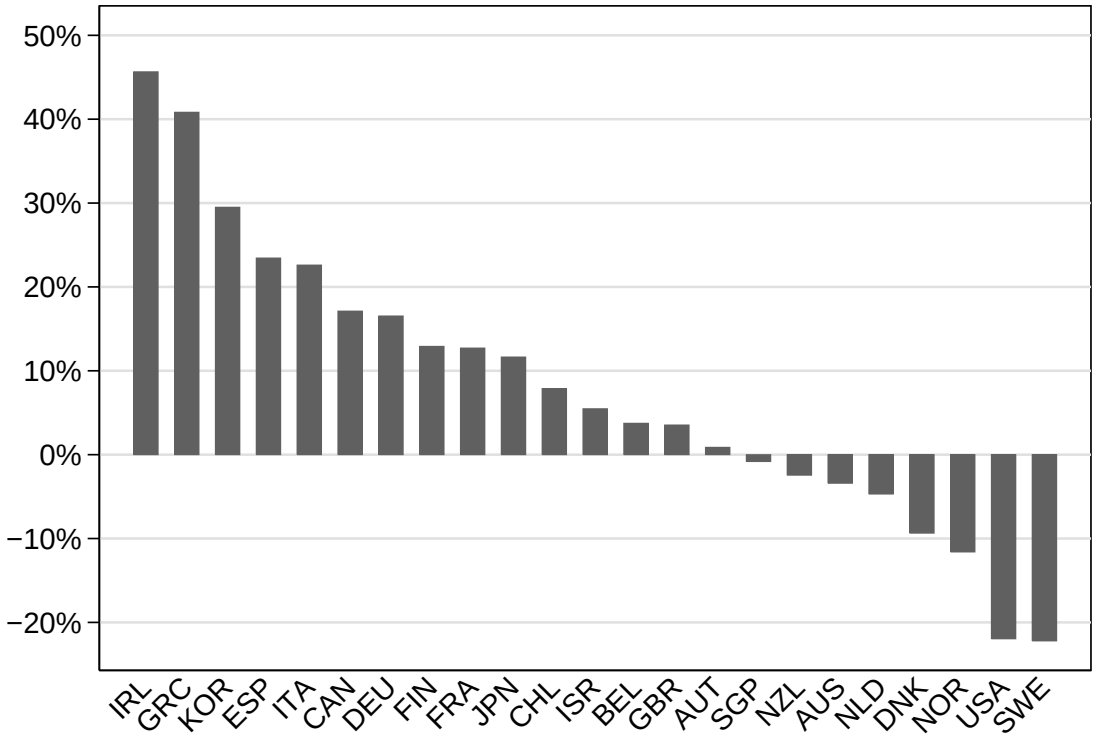
Note: The figure shows the coefficients on cognitive skills for various occupations. Dependent variable is student PISA test score in math (left graph), in reading (middle graph), and difference in standardized student test scores between math and reading (right graph). Skills in occupation refer to numeracy in left figure, to literacy in middle figure, and to the difference between numeracy and literacy in the right figure. Skills in occupation are z-standardized across countries. In the left graph, control variables are the same as in Column 3 of Table 5; in the middle graph, control variables are the same as in Column 6 of Table 5; in the right graph, control variables are the same as in Column 3 of Table 4. Occupations: "Teachers": teaching professionals; "Managers": administrative and commercial managers, production and specialised services managers, and hospitality, retail and other services managers; "Scientists&Engineers": science and engineering professionals and associate professionals; "Health workers": health professionals and associate professionals; "Business professionals": business and administration professionals; "Business associates": business and administration associate professionals; "Legal workers": legal, social and cultural professionals and associate professionals; "Clerks": general and keyboard clerks, customer services clerks,and numerical and material recording clerks; "Service workers": personal service workers; "Sales workers": sales workers; "Care workers": personal care workers; "Agricultural workers": skilled agricultural, forestry and fishery workers; "Craft workers": craft and related trades workers; "Operators": plant and machine operators, and assemblers; "Elementary workers": elementary occupations. Occupations are ranked by ISCO code (teachers placed first). The vertical dashed lines indicate the estimate on teacher cognitive skills. Asterisks next to the coefficient indicate the significance level (robust standard errors, adjusted for clustering at the country level): * p<0.10, ** p<0.05, *** p<0.01. *Data sources*: PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Figure 6: Placebo Tests Using Cognitive Skills of Matched Teacher Twins (OLS and Student Fixed Effects)



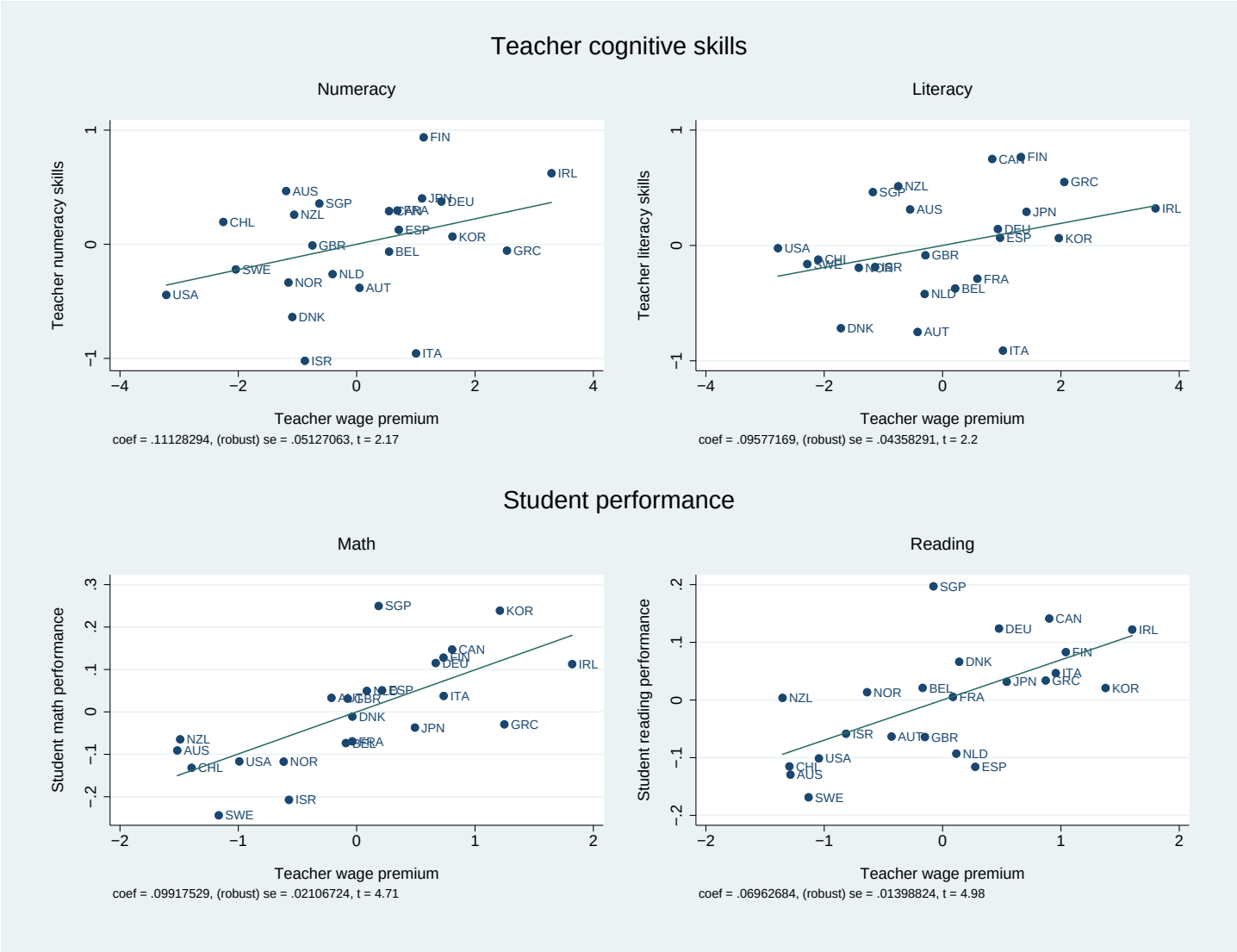
Note: The figure shows histograms of the coefficients on cognitive skills for 100 random samples of adults with the same sample size and age, gender, and education distribution as the teacher sample in the country. Dependent variable is student PISA test score in math (top graph), in reading (middle graph), and difference in standardized student test scores between math and reading (bottom graph). Skills refer to numeracy in top figure, to literacy in middle figure, and to the difference between numeracy and literacy in the bottom figure. Skills are z-standardized across countries. In the top graph, control variables are the same as in Column 3 of Table 5; in the middle graph, control variables are the same as in Column 6 of Table 5; in the bottom graph, control variables are the same as in Column 3 of Table 4. The vertical dashed lines indicate the estimate on teacher cognitive skills. *Data sources*: PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Figure 7: Teacher Wage Premiums around the World



Notes: Bars indicate the percentage difference in gross hourly earnings of teachers with a college degree relative to all nonteacher college graduates in a country. Estimates condition on gender, a quadratic polynomial in potential work experience (age − years of schooling − 6), and numeracy and literacy skills. Post-communist countries and Turkey are excluded (explanations see text). *Data sources:* PIAAC 2011/12 and 2014/15.

Figure 8: Teacher Wage Premiums and Teacher Cognitive Skills/Student Performance



Note: Dependent variable is standardized teacher cognitive skills (upper panel) and standardized student PISA test scores (lower panel), respectively. Upper panel shows added-variable plots that control for country-specific numeracy skills (left graph) and literacy skills (right graph) of all college graduates (without teachers); lower panel additionally controls for all variables included in the baseline specification in Table 2 (left graph: Column 3 of Table 2; right graph: Column 6 of Table 2). Teacher wage premiums are the percentage difference in gross hourly earnings of teachers with a college degree relative to all nonteacher college graduates in a country, conditional on gender, quadratic polynomial in potential work experience, and numeracy and literacy skills; divided by 10 (see also Figure 7 and Table EA-11). Post-communist countries and Turkey are excluded (explanations see text). *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table 1: Teacher Cognitive Skills by Country

	Pooled	Australia	Austria	Belgium	Canada	Chile	Czech R.	Denmark	Estonia	Finland	France
Numeracy	292	300	300	308	292	262	305	295	285	317	302
Literacy	295	312	292	303	307	263	300	288	294	322	296
Domain difference	-3	-12	8	5	-15	-1	5	7	-9	-5	6
Numeracy percentile	68	71	69	68	67	81	73	56	60	73	80
Literacy percentile	71	75	70	71	72	79	77	60	69	74	77
Observations	6,402	248	188	215	834	106	141	413	188	221	163
	Germany	Greece	Ireland	Israel	Italy	Japan	Korea	Lithuania	Netherl.	New Zealand	Norway
Numeracy	308	282	295	270	273	311	287	285	304	297	302
Literacy	301	286	300	281	279	319	296	282	308	310	304
Domain difference	7	-5	-4	-12	-5	-8	-9	3	-4	-12	-2
Numeracy percentile	72	74	75	57	67	70	72	66	63	64	65
Literacy percentile	74	75	74	62	73	67	74	64	67	71	68
Observations	127	150	180	250	124	147	217	133	197	198	279
		Poland	Russia	Singapore	Slovak R.	Slovenia	Spain	Sweden	Turkey	U.K.	U.S.
Numeracy		277	273	303	294	293	283	306	264	289	284
Literacy		293	283	300	290	288	290	307	261	299	301
Domain difference		-16	-10	3	4	5	-7	-1	3	-10	-17
Numeracy percentile		64	53	72	66	70	75	62	80	65	70
Literacy percentile		73	54	76	60	69	80	65	78	67	71
Observations		199	137	193	133	121	183	147	128	310	132

Notes: Teacher cognitive skills are country-specific median cognitive skills of primary school teachers, secondary school teachers, and "other" teachers (including, e.g., special education teachers and language teachers). Because occupation in Australia and Finland is reported only at the two-digit level, teachers in these countries include all "teaching professionals" (ISCO-08 code 23), i.e., additionally include pre-kindergarten teachers and university professors. All skill measures are rounded to the nearest integer. Percentile refers to the position of median cognitive skills of teachers in the cognitive skill distribution of all adults aged 25–65 excluding teachers. Individuals are weighted with PIAAC final sample weights. Observations refer to the number of teachers used to construct country-specific teacher skills. *Data sources:* PIAAC 2011/12 and 2014/15.

Table 2: Student Performance and Teacher Cognitive Skills (OLS)

	Student Math Performance			Student Reading Performance		
	(1)	(2)	(3)	(4)	(5)	(6)
Teacher cognitive skills	0.209***	0.173***	0.145***	0.178***	0.102***	0.092***
	(0.038)	(0.031)	(0.032)	(0.020)	(0.020)	(0.022)
Parent cognitive skills			0.044**			0.015
			(0.017)			(0.016)
Student characteristics		X	X		X	X
Parent characteristics		X	X		X	X
School characteristics		X	X		X	X
Country characteristics		X	X		X	X
Students	490,818	490,818	490,818	490,818	490,818	490,818
Countries	31	31	31	31	31	31
Adj. R2	0.04	0.29	0.29	0.03	0.30	0.30

Notes: Least squares regressions weighted by students' inverse sampling probability, giving each country the same weight. Dependent variable: student PISA test score in math (Columns 1–3) and in reading (Columns 4–6), respectively. Student test scores are z-standardized at the individual level across countries. Country-level teacher cognitive skills refer to numeracy in Columns 1–3 and to literacy in Columns 4–6. Teacher skills are z-standardized across countries. Parent cognitive skills are computed as the maximum of mother's and father's skills in numeracy (Columns 1–3) or literacy (Columns 4–6). Parent cognitive skills are standardized using teacher cognitive skills as "numeraire" scale. Student characteristics are age, gender, migrant status (first-generation or second-generation), and language spoken at home. Parent characteristics include parents' educational degree, number of books at home, and occupation. School characteristics include school location, number of students per school, three autonomy measures, as well as shortage of qualified teachers and weekly instructional time in math classes (Columns 1–3) or language classes (Columns 4–6). Country characteristics are expenditures per student and school starting age (Table EA-6 reports results for all control variables). All regressions include controls for respective imputation dummies and a dummy indicating the PISA wave. Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table 3: Simulation Analysis: Raising Teacher Cognitive Skills to Finnish Level

	Teacher Numeracy Skills		Teacher Literacy Skills	
	Difference from Finnish teachers (in PIAAC points)	Student perf. increase (in % of internat. SD)	Difference from Finnish teachers (in PIAAC points)	Student perf. increase (in % of internat. SD)
	(1)	(2)	(3)	(4)
Australia	17	17.8	10	6.7
Austria	17	17.2	30	19.6
Belgium	9	9.2	19	12.5
Canada	25	25.3	15	9.7
Chile	55	56.6	59	38.9
Czech Republic	12	12.4	22	14.8
Denmark	22	22.7	33	22.2
Estonia	32	33.1	28	18.8
France	16	16.0	26	17.2
Germany	9	9.1	21	13.6
Greece	36	36.4	36	23.7
Ireland	22	22.3	22	14.6
Israel	47	48.4	40	26.9
Italy	44	44.8	43	28.8
Japan	6	6.2	3	2.0
Korea	31	31.2	26	17.2
Lithuania	32	32.5	40	26.4
Netherlands	14	13.8	14	9.5
New Zealand	20	20.2	12	8.0
Norway	15	15.8	18	11.8
Poland	40	40.7	29	19.1
Russia	44	45.2	39	26.0
Singapore	14	14.6	22	14.7
Slovak Republic	23	23.3	32	21.2
Slovenia	25	25.1	34	22.6
Spain	34	35.1	32	21.2
Sweden	11	11.4	14	9.5
Turkey	53	54.4	60	40.1
U.K.	28	28.7	22	14.9
U.S.	33	33.7	21	13.8

Notes: This table shows by how much student performance would increase if teacher skills in numeracy and literacy, respectively, were at the levels in Finland (i.e., the country with highest teacher skills in both numeracy and literacy). Estimations are based on Columns 3 and 6 of Table 2. Columns 1 and 3 show difference in teacher skills to Finland. *Data sources:* PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table 4: Student Performance and Teacher Cognitive Skills (Student Fixed Effects)

Dependent Variable: Student Performance Difference: Math – Reading			
	(1)	(2)	(3)
Teacher skills: numeracy – literacy	0.105***	0.117***	0.106**
	(0.037)	(0.035)	(0.049)
Parent skills: numeracy – literacy			0.016
			(0.035)
Instruction time: math – reading		0.058**	0.058**
		(0.026)	(0.026)
Shortage teachers: math – reading		–0.012	–0.012
		(0.012)	(0.012)
Students	490,818	490,818	490,818
Countries	31	31	31
Adj. R2	0.01	0.02	0.02

Notes: Dependent variable: difference in standardized student test scores between math and reading. All regressions include controls for respective imputation dummies and for the PISA wave. Specifications give equal weight to each country. Robust standard errors, adjusted for clustering at country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table 5: Impact of Country-Level Adult Cognitive Skills on Student Performance (OLS)

	Student Math Performance			Student Reading Performance		
	(1)	(2)	(3)	(4)	(5)	(6)
Teacher cognitive skills	0.145***	0.134***	0.117**	0.092***	0.144***	0.148***
	(0.032)	(0.048)	(0.051)	(0.022)	(0.039)	(0.044)
Parent cognitive skills	0.044**	0.039***	0.033**	0.015	0.038**	0.035**
	(0.017)	(0.012)	(0.012)	(0.016)	(0.015)	(0.015)
Parent cognitive skills (country level)		0.014			–0.061*	
		(0.035)			(0.035)	
Adult cognitive skills (country level)			0.036			–0.064
			(0.040)			(0.041)
Student characteristics	X	X	X	X	X	X
Parent characteristics	X	X	X	X	X	X
School characteristics	X	X	X	X	X	X
Country characteristics	X	X	X	X	X	X
Students	490,818	490,818	490,818	490,818	490,818	490,818
Countries	31	31	31	31	31	31
Adj. R2	0.29	0.29	0.29	0.30	0.30	0.30

Notes: Dependent variable: standardized student PISA test score in math (Columns 1–3) and reading (Columns 4–6), respectively. All cognitive skill measures in Columns 1–3 (4–6) refer to numeracy (literacy) unless noted otherwise. Columns 1 and 4 report the baseline specification (see Columns 3 and 6 of Table 2). In Columns 2 and 5, we add the country-specific median cognitive skill level of PIAAC respondents aged 35–59 with children. In Columns 3 and 6, we add the median cognitive skill level of all PIAAC respondents aged 25–65. Student, parent, school, and country characteristics are the same as in the baseline least squares models (see Table 2). All regressions include controls for imputation dummies and the PISA wave. Specifications give equal weight to each country. Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table 6: Impact of Female Share in High-Skilled Occupations on Teacher Cognitive Skills

Dependent Variable: Teacher Cognitive Skills				
	Numeracy		Literacy	
	(1)	(2)	(3)	(4)
Share: female teachers/females in high-skilled occ. (×10)	0.371***	0.361***	0.243**	0.238*
	(0.120)	(0.123)	(0.124)	(0.126)
Numeracy skills of college graduates (w/o teachers)		0.271		
		(0.253)		
Literacy skills of college graduates (w/o teachers)				0.405*
				(0.224)
Country fixed effects	X	X	X	X
Cohort fixed effects	X	X	X	X
Observations	69	69	69	69
Countries	23	23	23	23

Notes: Dependent variable: teacher skills in numeracy (Columns 1–2) and literacy (Columns 3–4). Teacher cognitive skills are standardized using the standard deviation from the full sample (31 countries) as "numeraire" scale, such that magnitudes are comparable to the main analysis; cognitive skills of college graduates are standardized similarly. *Share: female teachers/females in high-skilled occ.* is the share of female teachers in a country-cohort cell over all females working in high-skilled occupations. Each cohort covers 15 adjacent birth years. Occupations are classified as high-skilled applying the following procedure in PIAAC: First, for each two-digit occupation in each country, we calculate average years of schooling of persons currently working in these occupations. Second, ranking occupations by average schooling level and starting from the occupation with the highest level, we define occupations as high-skilled until males working in these occupations comprise 25 percent of all males currently working in that country. Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. Post-communist countries and Turkey are excluded (explanations see text). *Data sources:* PIAAC 2011/12 and 2014/15.

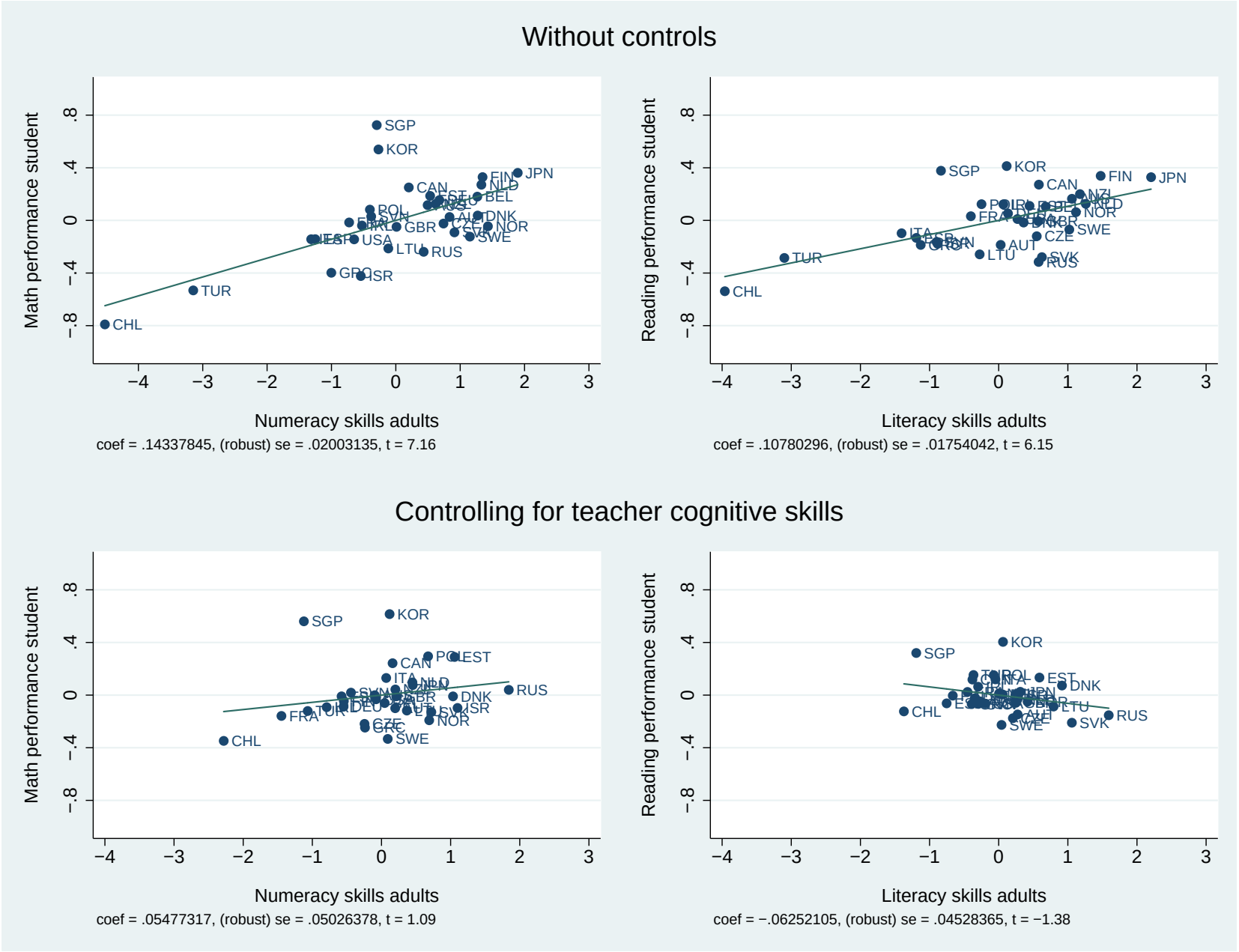
Table 7: Relationship of Teacher Wage Premiums to Teacher Cognitive Skills

Dependent Variable: Teacher Cognitive Skills		
	Numeracy	Literacy
	(1)	(2)
Teacher wage premium (/10)	0.113**	0.097**
	(0.052)	(0.044)
Numeracy skills of college graduates (w/o teachers)	0.943***	
	(0.112)	
Literacy skills of college graduates (w/o teachers)		0.918***
		(0.070)
Countries	23	23
Adj. R2	0.77	0.78

Notes: Dependent variable: teacher skills in numeracy (Column 1) and literacy (Column 2). *Teacher wage premium (/10)* is the percentage difference in gross hourly earnings of teachers with a college degree relative to all college graduates in a country, conditional on gender, quadratic polynomial in potential work experience, and numeracy and literacy skills; divided by 10. Robust standard errors in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. Post-communist countries and Turkey are excluded (explanations see text). *Data sources:* PIAAC 2011/12 and 2014/15.

Appendix

Figure A-1: Student Performance and Adult Cognitive Skills



Note: The two graphs in the top panel do not include any controls. The two graphs in the bottom panel are added-variable plots that control for country-level teacher skills in numeracy and literacy, respectively. *Data sources*: PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table A-1: Where Teachers Need to Be Drawn From to Get to the Skill Level of Finnish Teachers?

	Numeracy			Literacy		
	Current position teachers	Position to reach Finnish teachers	Difference	Current position teachers	Position to reach Finnish teachers	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Australia	55	69	14	56	66	10
Austria	44	59	15	45	76	31
Belgium	48	56	8	47	68	21
Canada	55	73	18	62	74	12
Chile	60	92	32	59	93	34
Czech R.	45	55	10	46	72	26
Denmark	42	61	19	44	78	34
Estonia	45	75	30	53	79	26
Finland	59	59	0	60	60	0
France	54	69	15	50	77	27
Germany	55	61	6	56	76	20
Greece	52	80	28	60	87	27
Ireland	58	77	19	57	78	21
Israel	44	77	33	54	85	31
Italy	42	79	37	46	84	38
Japan	53	60	7	56	58	2
Korea	52	82	30	55	83	28
Lithuania	41	70	29	45	85	40
Netherlands	46	58	12	46	61	15
New Zealand	53	69	16	58	70	12
Norway	44	58	14	50	68	18
Poland	38	74	36	45	72	27
Russia	49	87	38	54	85	31
Singapore	55	69	14	60	81	21
Slovak R.	38	61	23	44	80	36
Slovenia	50	72	22	51	84	33
Spain	54	85	31	56	85	29
Sweden	47	55	8	50	63	13
Turkey	50	89	39	53	97	44
U.K.	51	73	22	54	74	20
U.S.	47	74	27	51	71	20

Notes: Position refers to percentile in cognitive skill distribution of college educated. *Data source:* PIAAC 2011/12 and 2014/15.

Table A-2: Student Performance and Teacher Cognitive Skills (OLS): Same-Subject and Cross-Subject Effects

	Panel A: Teacher Numeracy Skills			
	Student Math Performance		Student Reading Performance	
	(1)	(2)	(3)	(4)
Teacher numeracy skills	0.145***	0.117**	0.067**	0.069*
	(0.032)	(0.051)	(0.028)	(0.038)
Parent cognitive skills	0.044**	0.033**	0.032	0.034*
	(0.017)	(0.012)	(0.022)	(0.017)
Cognitive skills of adults		0.036		–0.004
		(0.040)		(0.031)
Adj. R2	0.29	0.29	0.30	0.30

	Panel B: Teacher Literacy Skills			
	Student Math Performance		Student Reading Performance	
	(1)	(2)	(3)	(4)
Teacher literacy skills	0.116***	0.073*	0.092***	0.148***
	(0.029)	(0.042)	(0.022)	(0.044)
Parent cognitive skills	0.061***	0.045***	0.015	0.035**
	(0.015)	(0.015)	(0.016)	(0.015)
Cognitive skills of adults		0.051		–0.064
		(0.036)		(0.041)
Adj. R2	0.29	0.29	0.30	0.30

	Panel C: Teacher Numeracy and Literacy Skills			
	Student Math Performance		Student Reading Performance	
	(1)	(2)	(3)	(4)
Teacher numeracy skills	0.127*	0.117	0.013	0.011
	(0.069)	(0.073)	(0.052)	(0.049)
Teacher literacy skills	0.023	–0.000	0.082	0.139**
	(0.065)	(0.065)	(0.050)	(0.064)
Parent cognitive skills	0.043**	0.033**	0.015	0.034**
	(0.017)	(0.013)	(0.016)	(0.015)
Adult cognitive skills (country level)		0.037		–0.064
		(0.039)		(0.041)
Adj. R2	0.29	0.29	0.30	0.30
	Additional controls in Panels A–C			
Student characteristics	X	X	X	X
Parent characteristics	X	X	X	X
School characteristics	X	X	X	X
Country characteristics	X	X	X	X
Students	490,818	490,818	490,818	490,818
Countries	31	31	31	31

Notes: Least squares regressions weighted by students' inverse sampling probability, giving each country the same weight. Dependent variable: student PISA test score in math (Columns 1–2) and in reading (Columns 3–4), respectively. Student test scores are z-standardized at the individual level across countries. Teacher skills are z-standardized across countries. Parent skills and country-level adult skills refer to numeracy in Columns 1–2 and to literacy in Columns 3–4. Parent skills and country-level adult skills use teacher skills as "numeraire" scale. Control variables are the same as in the baseline least squares models (see Table 2). Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table A-3: Falsification Check Using Teacher ICT Skills (OLS)

	Student Math Performance			Student Reading Performance		
	(1)	(2)	(3)	(4)	(5)	(6)
Teacher ICT skills	0.081*	0.057	0.053	0.040	0.041	0.037
	(0.045)	(0.048)	(0.047)	(0.033)	(0.032)	(0.033)
Parent cognitive skills	0.076***		0.041**	0.041*		0.034*
	(0.018)		(0.015)	(0.021)		(0.019)
Adult cognitive skills (country level)		0.099***	0.077**		0.032	0.012
		(0.027)	(0.029)		(0.027)	(0.028)
Student characteristics	X	X	X	X	X	X
Parent characteristics	X	X	X	X	X	X
School characteristics	X	X	X	X	X	X
Country characteristics	X	X	X	X	X	X
Students	368,729	368,729	368,729	368,729	368,729	368,729
Countries	28	28	28	28	28	28
Adj. R2	0.28	0.28	0.29	0.30	0.30	0.30

Notes: Dependent variable: student PISA test score in math (Columns 1–3) and in reading (Columns 4–6), respectively. Student test scores are z-standardized at the individual level across countries. ICT skills were not tested in France, Italy, and Spain. Parent cognitive skills are computed as the maximum of mother's and father's skills in numeracy (Columns 1–3) or literacy (Columns 4–6). Country-level adult skills refer to numeracy in Columns 2–3 and to literacy in Columns 5–6. Parent skills and country-level adult skills use teacher skills (either in numeracy or in literacy) as "numeraire" scale. Control variables are the same as in the baseline least squares models (see Table 2). Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Electronic Appendix—Not for Publication

Table EA-1: Summary Statistics for Parent Cognitive Skills

	Pooled	Australia	Austria	Belgium	Canada	Chile	Czech R.	Denmark	Estonia	Finland	France
						Numeracy					
Mean	278	287	291	301	282	223	276	293	276	299	275
Std. Dev.	29	21	22	22	20	30	27	21	16	18	26
Max – Min	115	128	140	108	120	139	109	141	87	102	132
						Literacy					
Mean	275	293	279	289	284	226	270	278	272	297	272
Std. Dev.	26	19	20	20	18	23	24	20	16	17	21
Max – Min	101	113	111	96	116	105	98	148	95	101	106
Observations	83,492	3,137	2,231	2,251	11,933	2,165	2,105	3,352	3,463	2,252	3,086

	Germany	Greece	Ireland	Israel	Italy	Japan	Korea	Lithuania	Netherl.	New Zealand	Norway
						Numeracy					
Mean	289	273	275	267	267	308	276	277	295	284	297
Std. Dev.	21	19	22	25	19	14	17	20	22	22	23
Max – Min	126	77	96	132	104	50	85	65	120	134	192
						Literacy					
Mean	279	268	280	260	264	307	281	271	293	288	290
Std. Dev.	19	16	18	23	16	12	15	13	21	19	19
Max – Min	109	75	86	117	86	44	76	46	109	109	162
Observations	2,293	2,128	2,371	1,882	1,789	2,103	3,361	2,364	2,276	2,504	2,228

		Poland	Russia	Singapore	Slovak R.	Slovenia	Spain	Sweden	Turkey	U.K.	U.S.
						Numeracy					
Mean		264	271	261	281	268	265	295	240	281	267
Std. Dev.		19	8	39	23	24	22	25	27	20	32
Max – Min		103	32	149	139	149	94	174	100	109	135
						Literacy					
Mean		267	277	253	275	261	266	290	237	285	277
Std. Dev.		19	9	31	17	22	21	23	19	18	27
Max – Min		92	35	116	129	120	87	156	69	95	122
Observations		1,793	1,074	2,119	2,442	2,435	2,614	1,864	2,319	3,578	1,980

Notes: Summary statistics of parents' cognitive skills (average skill of mother and father) based on actual parents of PISA students. See text for computation of parent cognitive skills. *Max-Min* indicates the difference between the maximum and minimum parent cognitive skills within a country. *Observations* refer to the number of adults in the PIAAC samples used for computing parents' skills. *Data sources:* PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table EA-2: Summary Statistics for Student Performance and Student Characteristics

	Pooled	Australia	Austria	Belgium	Canada	Chile	Czech R.	Denmark	Estonia	Finland	France
Math performance	498	509	500	515	522	422	496	502	516	530	496
	(97)	(95)	(94)	(103)	(88)	(80)	(94)	(84)	(81)	(85)	(100)
Reading performance	497	513	480	508	524	445	486	496	508	530	501
	(97)	(98)	(96)	(102)	(91)	(81)	(91)	(84)	(82)	(91)	(108)
Age (in years)	15.8	15.8	15.8	15.8	15.8	15.8	15.8	15.7	15.8	15.7	15.9
Female	0.49	0.50	0.51	0.49	0.50	0.50	0.48	0.50	0.49	0.49	0.51
First-gen. migrant	0.06	0.12	0.06	0.09	0.13	0.01	0.02	0.04	0.01	0.02	0.05
Second-gen. migrant	0.05	0.12	0.11	0.08	0.15	0.00	0.01	0.06	0.07	0.01	0.10
Other language	0.09	0.09	0.11	0.22	0.16	0.01	0.02	0.05	0.04	0.04	0.08
Observations	490,818	28,732	11,345	17,098	44,751	12,525	11,391	13,405	9,506	14,639	8,911
	Germany	Greece	Ireland	Israel	Italy	Japan	Korea	Lithuania	Netherl.	New Zealand	Norway
Math performance	513	459	494	457	484	533	550	477	524	510	494
	(97)	(89)	(86)	(105)	(93)	(94)	(94)	(88)	(90)	(99)	(88)
Reading performance	503	479	509	480	488	529	537	473	510	517	503
	(93)	(97)	(92)	(113)	(96)	(100)	(83)	(87)	(91)	(104)	(96)
Age (in years)	15.8	15.7	15.7	15.7	15.7	15.8	15.7	15.8	15.7	15.8	15.8
Female	0.49	0.51	0.49	0.51	0.48	0.48	0.47	0.49	0.50	0.49	0.49
First-gen. migrant	0.05	0.07	0.12	0.07	0.06	0.00	0.00	0.01	0.04	0.19	0.05
Second-gen. migrant	0.11	0.04	0.02	0.12	0.02	0.00	0.00	0.01	0.08	0.09	0.04
Other language	0.09	0.05	0.05	0.11	0.14	0.00	0.00	0.04	0.06	0.15	0.07
Observations	9,980	10,094	8,953	10,816	61,978	12,439	10,022	9,146	9,220	8,934	9,346
		Poland	Russia	Singapore	Slovak R.	Slovenia	Spain	Sweden	Turkey	U.K.	U.S.
Math performance		506	475	568	489	501	484	486	447	493	484
		(90	(86)	(105)	(99)	(93)	(89)	(93)	(92)	(91)	(90)
Reading performance		509	467	534	470	482	485	491	470	497	498
		(89)	(90)	(100)	(98)	(91)	(90)	(103)	(84)	(96)	(94)
Age (in years)		15.7	15.8	15.8	15.8	15.7	15.9	15.7	15.8	15.7	15.8
Female		0.51	0.50	0.49	0.49	0.49	0.49	0.49	0.49	0.51	0.49
First-gen. migrant		0.00	0.05	0.12	0.01	0.03	0.10	0.06	0.00	0.07	0.07
Second-gen. migrant		0.00	0.07	0.05	0.00	0.06	0.01	0.08	0.01	0.05	0.13
Other language		0.01	0.09	0.57	0.06	0.06	0.18	0.09	0.05	0.07	0.14
Observations		9,524	10,539	10,829	9,233	12,066	51,200	9,303	9,844	24,838	10,211

Notes: Means and standard deviations (in parentheses) reported. *Other language* indicates a student who speaks a foreign language at home. *Observations* refer to the number of students in both PISA cycles. Statistics are based on student-level observations weighted with inverse sampling probabilities, giving each PISA cycle the same total weight. *Data sources:* PISA 2009 and 2012.

Table EA-3: Summary Statistics for Parent Characteristics

	Pooled	Australia	Austria	Belgium	Canada	Chile	Czech R.	Denmark	Estonia	Finland	France
Number of books at home											
0-10 books	0.13	0.09	0.13	0.16	0.10	0.23	0.10	0.13	0.07	0.07	0.16
11-25 books	0.16	0.12	0.16	0.17	0.14	0.29	0.14	0.16	0.14	0.12	0.17
26-100 books	0.32	0.30	0.31	0.29	0.31	0.31	0.35	0.32	0.31	0.34	0.29
101-200 books	0.18	0.21	0.17	0.17	0.21	0.10	0.19	0.18	0.21	0.22	0.17
201-500 books	0.14	0.18	0.14	0.13	0.16	0.05	0.15	0.14	0.17	0.18	0.13
More than 500 books	0.08	0.10	0.09	0.08	0.08	0.02	0.07	0.07	0.09	0.06	0.07
Highest educational degree											
ISCED 0	0.01	0.00	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.01
ISCED 1	0.02	0.01	0.01	0.02	0.01	0.03	0.00	0.01	0.00	0.01	0.01
ISCED 2	0.06	0.05	0.04	0.03	0.02	0.18	0.01	0.05	0.03	0.02	0.09
ISCED 3B,C	0.09	0.07	0.29	0.05	0.00	0.00	0.18	0.13	0.02	0.08	0.19
ISCED 3A,4	0.29	0.32	0.18	0.28	0.25	0.43	0.49	0.15	0.38	0.09	0.19
ISCED 5B	0.19	0.13	0.28	0.22	0.24	0.12	0.09	0.41	0.22	0.27	0.22
ISCED 5A,6	0.34	0.42	0.20	0.40	0.48	0.22	0.23	0.24	0.35	0.53	0.30
Highest occupational status											
Blue collar-low skilled	0.07	0.05	0.05	0.09	0.06	0.16	0.07	0.05	0.06	0.03	0.07
Blue collar-high skilled	0.11	0.08	0.14	0.10	0.07	0.17	0.13	0.07	0.14	0.07	0.11
White collar-low skilled	0.24	0.17	0.26	0.23	0.21	0.28	0.27	0.25	0.23	0.20	0.26
White collar-high skilled	0.56	0.68	0.53	0.56	0.64	0.34	0.52	0.62	0.55	0.69	0.54

Table EA-3: Summary Statistics for Parent Characteristics (continued)

	Germany	Greece	Ireland	Israel	Italy	Japan	Korea	Lithuania	Netherl.	New Zealand	Norway
Number of books at home											
0-10 books	0.11	0.11	0.14	0.12	0.12	0.09	0.05	0.16	0.16	0.10	0.08
11-25 books	0.13	0.20	0.15	0.17	0.19	0.13	0.09	0.20	0.18	0.13	0.11
26-100 books	0.10	0.08	0.07	0.12	0.08	0.09	0.10	0.05	0.07	0.09	0.11
101-200 books	0.29	0.32	0.30	0.30	0.30	0.35	0.29	0.33	0.30	0.31	0.30
201-500 books	0.20	0.17	0.19	0.17	0.18	0.19	0.23	0.15	0.15	0.21	0.22
More than 500 books	0.17	0.12	0.15	0.13	0.13	0.15	0.24	0.10	0.13	0.17	0.19
Highest educational degree											
ISCED 0	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00
ISCED 1	0.00	0.03	0.02	0.01	0.01	0.00	0.01	0.00	0.02	0.01	0.00
ISCED 2	0.15	0.09	0.07	0.03	0.21	0.02	0.04	0.01	0.04	0.06	0.02
ISCED 3B,C	0.12	0.02	0.02	0.09	0.06	0.06	0.07	0.01	0.00	0.16	0.03
ISCED 3A,4	0.23	0.34	0.35	0.26	0.37	0.30	0.34	0.37	0.32	0.25	0.25
ISCED 5B	0.18	0.14	0.18	0.16	0.07	0.15	0.06	0.19	0.39	0.15	0.39
ISCED 5A,6	0.30	0.37	0.35	0.44	0.28	0.47	0.48	0.41	0.21	0.37	0.30
Highest occupational status											
Blue collar-low skilled	0.06	0.08	0.05	0.07	0.07	0.07	0.04	0.08	0.04	0.07	0.03
Blue collar-high skilled	0.10	0.14	0.09	0.06	0.17	0.08	0.06	0.18	0.06	0.07	0.04
White collar-low skilled	0.29	0.24	0.26	0.15	0.28	0.36	0.29	0.22	0.20	0.18	0.16
White collar-high skilled	0.53	0.51	0.58	0.68	0.45	0.48	0.59	0.49	0.68	0.66	0.75

Table EA-3: Summary Statistics for Parent Characteristics (continued)

	Poland	Russia	Singapore	Slovak R.	Slovenia	Spain	Sweden	Turkey	U.K.	U.S.
Number of books at home										
0-10 books	0.11	0.09	0.11	0.15	0.14	0.09	0.09	0.26	0.14	0.21
11-25 books	0.20	0.19	0.19	0.17	0.20	0.15	0.11	0.26	0.16	0.18
26-100 books	0.07	0.08	0.05	0.05	0.06	0.09	0.11	0.03	0.08	0.05
101-300 books	0.34	0.34	0.36	0.37	0.35	0.32	0.30	0.28	0.29	0.29
301-500 books	0.17	0.17	0.17	0.17	0.15	0.21	0.20	0.11	0.18	0.15
More than 500 books	0.11	0.13	0.12	0.10	0.10	0.15	0.19	0.06	0.15	0.11
Highest educational degree										
ISCED 0	0.00	0.00	0.01	0.00	0.00	0.02	0.01	0.04	0.00	0.01
ISCED 1	0.00	0.00	0.05	0.00	0.00	0.07	0.01	0.32	0.00	0.02
ISCED 2	0.01	0.05	0.02	0.04	0.18	0.04	0.24	0.03	0.05	
ISCED 3B,C	0.39	0.01	0.00	0.14	0.35	0.02	0.07	0.02	0.20	0.00
ISCED 3A,4	0.33	0.08	0.44	0.54	0.19	0.25	0.18	0.17	0.18	0.34
ISCED 5B	0.00	0.44	0.19	0.06	0.16	0.14	0.21	0.08	0.23	0.15
ISCED 5A,6	0.24	0.46	0.27	0.23	0.25	0.33	0.48	0.14	0.36	0.43
Highest occupational status										
Blue collar-low skilled	0.07	0.06	0.06	0.11	0.07	0.09	0.05	0.14	0.05	0.07
Blue collar-high skilled	0.27	0.11	0.04	0.16	0.14	0.18	0.05	0.25	0.05	0.06
White collar-low skilled	0.23	0.26	0.21	0.31	0.24	0.29	0.24	0.25	0.26	0.21
White collar-high skilled	0.43	0.54	0.67	0.40	0.53	0.43	0.65	0.28	0.62	0.64

Notes: Shares reported. Statistics are based on student-level observations weighted with inverse sampling probabilities, giving each PISA cycle the same total weight. *Highest educational degree* includes the following categories: *ISCED 0:* no educational degree; *ISCED 1:* primary education; *ISCED 2:* lower secondary; *ISCED 3B,C:* vocational/pre-vocational upper secondary; *ISCED 3A,4:* upper secondary or non-tertiary post-secondary; *ISCED 5B:* vocational tertiary; and *ISCED 5A,6:* theoretically oriented tertiary and post-graduate. *Data sources:* PISA 2009 and 2012.

Table EA-4: Summary Statistics for School Characteristics

	Pooled	Australia	Austria	Belgium	Canada	Chile	Czech R.	Denmark	Estonia	Finland	France
Instructional time math	3.6	4.0	2.6	3.5	5.3	5.8	3.1	3.7	3.7	2.9	3.5
Instructional time reading	3.6	3.9	2.4	3.6	5.4	5.7	3.0	5.2	3.3	2.5	3.7
Shortage math teachers	1.52	1.89	1.33	1.92	1.44	2.05	1.25	1.23	1.45	1.16	1.35
Shortage language teachers	1.42	1.53	1.36	1.54	1.26	1.82	1.12	1.17	1.30	1.10	1.36
Private school	0.19	0.41	0.11	0.69	0.08	0.61	0.06	0.24	0.04	0.04	0.20
Students per school	735	981	559	718	1032	1013	450	480	557	429	821
Content autonomy	0.64	0.71	0.58	0.56	0.37	0.67	0.88	0.68	0.77	0.64	0.64
Personnel autonomy	0.42	0.39	0.08	0.38	0.30	0.63	0.88	0.58	0.54	0.24	0.06
Budget autonomy	0.82	0.93	0.86	0.69	0.75	0.78	0.79	0.96	0.84	0.92	0.97
	Germany	Greece	Ireland	Israel	Italy	Japan	Korea	Lithuania	Netherl.	New Zealand	Norway
Instructional time math	3.3	3.4	3.1	4.3	3.8	3.9	3.6	2.9	2.8	4.0	3.2
Instructional time reading	3.1	3.0	3.0	3.4	4.7	3.5	3.5	3.4	2.8	4.1	3.8
Shortage math teachers	1.78	1.13	1.40	1.90	1.69	1.27	1.57	1.14	2.10	1.72	1.73
Shortage language teachers	1.46	1.20	1.16	1.96	1.64	1.21	1.57	1.14	1.74	1.40	1.70
Private school	0.06	0.06	0.60	0.09	0.06	0.30	0.42	0.01	0.67	0.06	0.02
Students per school	702	283	593	770	752	750	1116	593	1023	1178	340
Content autonomy	0.63	0.04	0.69	0.53	0.72	0.92	0.89	0.80	0.93	0.88	0.49
Personnel autonomy	0.15	0.03	0.34	0.39	0.05	0.32	0.23	0.65	0.89	0.55	0.42
Budget autonomy	0.88	0.84	0.87	0.69	0.84	0.90	0.85	0.59	0.99	0.99	0.88
		Poland	Russia	Singapore	Slovak R.	Slovenia	Spain	Sweden	Turkey	U.K.	U.S.
Instructional time math		3.4	3.6	5.4	3.0	2.7	3.5	3.1	2.9	3.7	4.3
Instructional time reading		3.7	3.1	4.3	3.0	2.9	3.4	3.0	3.6	3.8	4.4
Shortage math teachers		1.03	1.71	1.35	1.13	1.12	1.09	1.35	2.73	1.64	1.37
Shortage language teachers		1.01	1.63	1.95	1.10	1.06	1.08	1.19	2.64	1.38	1.20
Private school		0.03	0.00	0.02	0.09	0.03	0.33	0.12	0.01	0.26	0.08
Students per school		324	566	1367	480	462	701	420	890	1062	1381
Content autonomy		0.75	0.59	0.63	0.59	0.45	0.53	0.63	0.20	0.89	0.48
Personnel autonomy		0.46	0.65	0.10	0.70	0.51	0.18	0.72	0.02	0.75	0.66
Budget autonomy		0.26	0.58	0.89	0.72	0.79	0.94	0.93	0.77	0.96	0.76

Notes: Country means reported. Student-level information on *instructional time* (hours per week) is aggregated to the school level for both math and reading (see also Lavy (2015)). *Shortage math/language teachers* is based on the following school principal question: "Is your school's capacity to provide instruction hindered by any of the following issues? A lack of qualified mathematics/test language teachers" Possible answer categories are: not at all (1), very little (2), to some extent (3), a lot (4). School autonomy measures are binary. *Data sources:* PISA 2009 and 2012.

Table EA-5: Summary Statistics for Country Characteristics

	Pooled	Australia	Austria	Belgium	Canada	Chile	Czech R.	Denmark	Estonia	Finland	France
Expenditure per student	70.79	85.21	107.20	88.64	80.42	27.92	49.64	98.69	49.28	78.81	79.12
School starting age	6.12	5	6	6	5	6	6	7	7	7	6
Instruction practice math	0.61	0.66	0.57	0.56	0.70	0.67	0.62	0.64	0.59	0.58	0.59
Instruction practice reading	0.50	0.53	0.41	0.43	0.56	0.53	0.44	0.57	0.50	0.37	0.52
GDP per capita	35.34	41.43	43.24	39.78	40.45	18.80	27.87	41.93	23.06	38.99	36.13
Teacher gross hourly wage	18.9	21.4	19.6	23.6	26.6	14.2	9.4	22.9	9.1	22.6	21.1
Teacher performance pay	0.59	0	1	0	0	1	1	1	1	1	0
Central exit exams	0.70	1.0	0.0	0.0	0.7	0.0	0.5	1.0	1.0	1.0	1.0

	Germany	Greece	Ireland	Israel	Italy	Japan	Korea	Lithuania	Netherl.	New Zealand	Norway
Expenditure per student	72.05	53.29	84.52	55.17	80.86	83.70	65.07	41.20	87.71	59.64	112.43
School starting age	6	6	4	6	6	6	6	7	6	5	6
Instruction practice math	0.64	0.62	0.69	0.69	0.59	0.46	0.38	0.63	0.57	0.66	0.52
Instruction practice reading	0.44	0.49	0.51	0.40	0.49	0.44	0.34	0.58	0.37	0.53	0.37
GDP per capita	40.36	28.32	43.96	29.78	35.02	33.80	30.31	21.38	45.42	31.89	60.78
Teacher gross hourly wage	26.7	18.8	35.7	14.7	23.0	18.4	25.0	11.0	22.3	19.8	23.6
Teacher performance pay	0	0	0	0	0	0	0	.	1	1	1
Central exit exams	0.9	0.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

	Poland	Russia	Singapore	Slovak R.	Slovenia	Spain	Sweden	Turkey	U.K.	U.S.
Expenditure per student	48.80	13.29	78.15	42.68	84.84	78.15	89.29	16.26	91.46	110.86
School starting age	7	7	7	6	6	6	7	7	5	6
Instruction practice math	0.60	0.69	0.70	0.54	0.56	0.64	0.51	0.59	0.73	0.72
Instruction practice reading	0.59	0.80	0.47	0.47	0.56	0.44	0.42	0.64	0.54	0.61
GDP per capita	21.37	22.35	69.37	24.63	27.99	32.52	42.05	16.55	36.97	49.22
Teacher gross hourly wage	12.8	4.7	22.4	8.6	11.7	19.8	16.4	19.7	21.2	20.0
Teacher performance pay	1	1	.	1	1	0	1	1	1	1
Central exit exams	1.0	.	.	1.0	1.0	0.0	0.0	0.0	1.0	0.1

Notes: Expenditure per student and GDP per capita are expressed in 1,000 PPP-US-$. The *instruction practice* indicators are based on student information provided in PISA; in 2009 for language teachers and in 2012 for math teachers. See text for details on the construction of the instruction practice indicators. *Teacher performance pay* is a binary variable, taking the value 1 if salary adjustments are awarded to teachers with outstanding performance in teaching in a country; not available for Lithuania and Singapore. *Central exit exams* equals 1 if central exam examinations exist on the upper secondary level (ISCED 3) in a country, 0 otherwise; data are taken from Leschnig, Schwerdt, and Zigova (2016). Information on central exit exams is not available for the Russian Federation and Singapore. The remaining country characteristics come from OECD statistics. *Data sources:* Leschnig, Schwerdt, and Zigova (2016), OECD, PISA 2009 and 2012.

Table EA-6: Student Performance and Teacher Cognitive Skills from OLS Estimation: Results on Covariates not Reported in Table 2

Dependent variable: student performance	Math	Reading
Student characteristics		
Age	0.137***	0.137***
	(0.018)	(0.012)
Female	–0.145***	0.358***
	(0.011)	(0.015)
First-generation migrant	–0.107***	–0.103**
	(0.038)	(0.038)
Second-generation migrant	–0.086**	–0.021
	(0.035)	(0.034)
Other language at home	–0.056*	–0.179***
	(0.029)	(0.031)
Family background		
Books at home		
11-25 books	0.186***	0.226***
	(0.021)	(0.021)
26-100 books	0.420***	0.467***
	(0.033)	(0.034)
101-200 books	0.588***	0.647***
	(0.043)	(0.044)
201-500 books	0.776***	0.822***
	(0.049)	(0.053)
More than 500 books	0.775***	0.801***
	(0.053)	(0.059)
Parental education		
ISCED 1	0.175***	0.219***
	(0.042)	(0.042)
ISCED 2	0.090	0.137**
	(0.065)	(0.054)
ISCED 3B,C	0.254***	0.242***
	(0.069)	(0.060)
ISCED 3A, 4	0.249***	0.270***
	(0.062)	(0.055)
ISCED 5B	0.169*	0.244***
	(0.089)	(0.074)
ISCED 5A, 6	0.261***	0.330***
	(0.085)	(0.067)
Parental occupation		
Blue collar-high skilled	0.119***	0.097***
	(0.015)	(0.018)
White collar-low skilled	0.190***	0.184***
	(0.016)	(0.019)
White collar-high skilled	0.403***	0.405***
	(0.018)	(0.020)

(continued on next page)

Table EA-6 (continued)

Dependent variable: student performance	Math	Reading
School characteristics		
School location		
Small Town	–0.008	0.019
	(0.032)	(0.028)
Town	0.014	0.057
	(0.042)	(0.035)
City	0.014	0.079**
	(0.040)	(0.034)
Large City	0.080*	0.129***
	(0.045)	(0.043)
Private school	0.140***	0.159***
	(0.038)	(0.031)
No. students per school (in 1000)	0.281***	0.255***
	(0.062)	(0.052)
School autonomy		
Content autonomy	0.069	0.002
	(0.051)	(0.032)
Personnel autonomy	–0.148***	–0.167***
	(0.048)	(0.031)
Budget autonomy	0.020	0.048
	(0.039)	(0.036)
Shortage math teacher	–0.048***	
	(0.012)	
Shortage language teacher		–0.032**
		(0.013)
Weekly hours math classes	0.057**	
	(0.027)	
Weekly hours language classes		–0.001
		(0.018)
Country-level measures		
Educational expenditure per student	–0.000	0.000
	(0.001)	(0.001)
School starting age	0.139***	0.080*
	(0.049)	(0.041)
Students	490,818	490,818
Countries	31	31
Adj. R2	0.29	0.30

Notes: The table reports results on all further covariates of the ordinary least squares estimations with the full set of control variables, corresponding to Column 3 (math) and Column 6 (reading) in Table 2. Omitted categories of family background and school characteristics: *0-10 books*; *parents have no educational degree*; *blue collar-low skilled*; and *village*. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table EA-7: Student Performance and Teacher Cognitive Skills (Heterogeneity)

	Gender		Parental background		Migration background	
	Boys	Girls	High SES	Low SES	Natives	Migrants
Panel A: Student Math Performance						
Teacher cognitive skills	0.135***	0.155***	0.137***	0.144***	0.140***	0.107**
	(0.032)	(0.032)	(0.032)	(0.033)	(0.034)	(0.045)
Parent cognitive skills	0.046***	0.040**	0.079***	0.025	0.049**	0.061**
	(0.016)	(0.019)	(0.020)	(0.019)	(0.018)	(0.026)
Panel B: Student Reading Performance						
Teacher cognitive skills	0.081***	0.103***	0.068***	0.112***	0.082***	0.070*
	(0.021)	(0.025)	(0.024)	(0.024)	(0.023)	(0.038)
Parent cognitive skills	0.016	0.013	0.052**	0.004	0.022	0.017
	(0.015)	(0.018)	(0.025)	(0.015)	(0.018)	(0.023)
Students	246,649	244,169	250,954	239,864	424,419	24,232
Countries	31	31	31	31	31	30
Additional controls in Panels A + B						
Student characteristics	X	X	X	X	X	X
Parent characteristics	X	X	X	X	X	X
School characteristics	X	X	X	X	X	X
Country characteristics	X	X	X	X	X	X

Notes: Dependent variable: standardized student PISA test score in math (Panel A) and reading (Panel B), respectively. Parental background is measured by the PISA index of economic, social and cultural status (ESCS). This index captures a range of aspects of a student's family and home background that combines information on parents' education, occupations, and home possessions. Migrants refer to second-generation migrants. To account for the unequal distribution of migrants across countries, we re-weight regressions based on the sample of natives and migrants, respectively, giving equal weight to each country within each subsample. Korea has no second-generation migrants in the PISA sample and is therefore excluded. All cognitive skill measures in Panel A (Panel B) refer to numeracy (literacy). Student, parent, school, and country characteristics are the same as in the least squares models (see Table 2). All regressions include controls for respective imputation dummies and a dummy indicating the PISA wave. Specifications give equal weight to each country. Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table EA-8: Student Performance and Position of Teachers in Adult Cognitive Skill Distribution (OLS)

	Student Math Performance			Student Reading Performance		
	(1)	(2)	(3)	(4)	(5)	(6)
Position of teacher skills	0.015***	0.014***	0.013***	0.020***	0.014***	0.013***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Parent cognitive skills			0.029**			0.026*
			(0.012)			(0.014)
Adult cognitive skills (country level)	0.184***	0.159***	0.140***	0.149***	0.083***	0.067***
	(0.022)	(0.025)	(0.026)	(0.017)	(0.020)	(0.021)
Student characteristics		X	X		X	X
Parent characteristics		X	X		X	X
School characteristics		X	X		X	X
Country characteristics		X	X		X	X
Students	490,818	490,818	490,818	490,818	490,818	490,818
Countries	31	31	31	31	31	31
Adj. R2	0.04	0.29	0.29	0.03	0.30	0.30

Notes: Least squares regressions weighted by students' inverse sampling probability, giving each country the same weight. Dependent variable: student PISA test score in math (Columns 1–3) and in reading (Columns 4–6), respectively. Student test scores are z-standardized at the individual level across countries. *Position of teacher skills* is the country-specific percentile rank of teacher cognitive skills in the cognitive skill distribution all adults aged 25–65 years. Position of teacher skills, parent skills, and country-level adult skills refer to numeracy in Columns 1–3 and to literacy in Columns 4–6. Parent skills and country-level adult skills use teacher skills as "numeraire" scale. Control variables are the same as in the baseline least squares models (see Table 2). Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table EA-9: Student Performance and Teacher Cognitive Skills: Including Additional Country Controls

	Student Math Performance					Student Reading Performance				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Teacher cognitive skills	0.145***	0.147***	0.134***	0.108***	0.123***	0.092***	0.101***	0.080***	0.070***	0.084***
	(0.032)	(0.033)	(0.030)	(0.029)	(0.028)	(0.022)	(0.023)	(0.024)	(0.021)	(0.022)
Parent cognitive skills	0.044**	0.045**	0.047**	0.052***	0.053***	0.015	0.012	0.022	0.032*	0.026*
	(0.017)	(0.017)	(0.017)	(0.014)	(0.014)	(0.016)	(0.014)	(0.015)	(0.017)	(0.014)
Instructional practices		0.191			0.104		0.291			0.593***
		(0.191)			(0.164)		(0.211)			(0.156)
GDP per capita (1,000 PPP-$)			0.003		−0.008**			0.003		0.004*
			(0.004)		(0.003)			(0.003)		(0.002)
Central exit exams				0.167***	0.166***				0.095**	0.095***
				(0.056)	(0.051)				(0.036)	(0.034)
Student characteristics	X	X	X	X	X	X	X	X	X	X
Parent characteristics	X	X	X	X	X	X	X	X	X	X
School characteristics	X	X	X	X	X	X	X	X	X	X
Country characteristics	X	X	X	X	X	X	X	X	X	X
Students	490,818	490,818	490,818	469,450	469,450	490,818	490,818	490,818	469,450	469,450
Countries	31	31	31	29	29	31	31	31	29	29
Adj. R2	0.29	0.29	0.29	0.29	0.29	0.30	0.30	0.30	0.30	0.31

Notes: Dependent variable: standardized student PISA test score in math (Columns 1–5) and reading (Columns 6–10), respectively. Teacher and parent cognitive skills in Columns 1–5 (6–10) refer to numeracy (literacy). Columns 1 and 6 replicate the baseline models from Columns 3 and 6 in Table 2. Indicator for teacher instructional practices is based on the PISA data. See text for details on the construction of the instructional practices indicator. *Central exit exams* takes the value 1 if central exam examinations exist on the upper secondary level (ISCED 3) in a country; data are taken from Leschnig, Schwerdt, and Zigova (2016). Information on central exit exams is not available for the Russian Federation and Singapore. Student, parent, school, and country characteristics are the same as in the baseline least squares models (see Table 2). All regressions include controls for imputation dummies and the PISA wave. Specifications give equal weight to each country. Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, Leschnig, Schwerdt, and Zigova (2016), PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table EA-10: Student Performance and Teacher Cognitive Skills with Continental Fixed Effects and in Country Subsamples (OLS)

		Continent		w/o ex-communist	Large
	Baseline	Fixed effects	Europe only	& Turkey	teacher sample
	(1)	(2)	(3)	(4)	(5)
Panel A: Student Math Performance					
Teacher cognitive skills	0.145***	0.127***	0.104***	0.178***	0.171***
	(0.032)	(0.030)	(0.030)	(0.032)	(0.045)
Parent cognitive skills	0.044**	0.024**	0.038**	0.034**	0.004
	(0.017)	(0.012)	(0.014)	(0.014)	(0.010)
Panel B: Student Reading Performance					
Teacher cognitive skills	0.092***	0.088***	0.072**	0.102***	0.118***
	(0.022)	(0.025)	(0.029)	(0.028)	(0.020)
Parent cognitive skills	0.015	0.003	0.016	0.003	–0.019
	(0.016)	(0.012)	(0.015)	(0.018)	(0.018)
Students	490,818	490,818	352,375	409,569	312,163
Countries	31	31	23	23	19
Additional controls in Panels A + B					
Student characteristics	X	X	X	X	X
Parent characteristics	X	X	X	X	X
School characteristics	X	X	X	X	X
Country characteristics	X	X	X	X	X

Notes: Dependent variable: standardized student PISA test score in math (Panel A) and reading (Panel B). All skill measures in Panel A (Panel B) refer to numeracy (literacy). Column 1 replicates the baseline least squares models from Columns 3 and 6 in Table 2. In Column 2, we add continental fixed effects and in Column 3, the sample is restricted to only European countries. In Column 4, we exclude countries with a communist heritage and Turkey, while we keep only countries with at least 150 teacher observations in PIAAC in Column 5. Student, parent, school, and country characteristics are the same as in the baseline least squares models (see Table 2). All regressions include controls for imputation dummies and the PISA wave. Specifications give equal weight to each country. Robust standard errors, adjusted for clustering at the country level, in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* OECD, PIAAC 2011/12 and 2014/15, PISA 2009 and 2012.

Table EA-11: Teacher Wage Premiums around the World: Regression Output

	Australia	Austria	Belgium	Canada	Chile	Denmark	Finland	France	Germany	Greece	Ireland	Israel
Teacher	–0.034	0.009	0.037	0.171***	0.079	–0.094***	0.129***	0.127***	0.165***	0.408***	0.457***	0.055
	(0.027)	(0.034)	(0.025)	(0.025)	(0.091)	(0.016)	(0.021)	(0.032)	(0.042)	(0.064)	(0.040)	(0.052)
Numeracy	0.124***	0.015	0.058***	0.082***	0.174***	0.046***	0.103***	0.076***	0.062**	0.027	0.095***	0.133***
	(0.019)	(0.027)	(0.017)	(0.013)	(0.049)	(0.014)	(0.014)	(0.017)	(0.026)	(0.042)	(0.025)	(0.022)
Literacy	–0.013	0.105***	0.016	0.073***	–0.007	0.040***	–0.009	0.013	0.073 ***	0.014	0.037	0.050**
	(0.020)	(0.026)	(0.017)	(0.013)	(0.054)	(0.014)	(0.015)	(0.018)	(0.026)	(0.037)	(0.026)	(0.023)
Female	–0.120***	–0.113***	–0.036*	–0.113***	–0.210***	–0.116***	–0.162***	–0.061***	–0.168***	–0.071	0.004	–0.102***
	(0.021)	(0.030)	(0.019)	(0.016)	(0.060)	(0.016)	(0.017)	(0.020)	(0.029)	(0.048)	(0.029)	(0.035)
Pot. experience	0.036***	0.026***	0.026***	0.039***	0.035***	0.026***	0.024***	0.031***	0.046***	0.042***	0.052***	0.051***
	(0.003)	(0.005)	(0.003)	(0.003)	(0.009)	(0.003)	(0.003)	(0.003)	(0.006)	(0.009)	(0.005)	(0.006)
Pot. experience^2	–0.001***	–0.000**	–0.000***	–0.001***	–0.001***	–0.000***	–0.000***	–0.000***	–0.001***	–0.000*	–0.001***	–0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
		Italy	Japan	Korea	Netherl.	New Zeal.	Norway	Singapore	Spain	Sweden	U.K.	U.S.
Teacher		0.226***	0.116**	0.295***	–0.047	–0.025	–0.116***	–0.008	0.234***	–0.222***	0.035	–0.220***
		(0.055)	(0.049)	(0.052)	(0.030)	(0.033)	(0.016)	(0.043)	(0.035)	(0.021)	(0.039)	(0.039)
Numeracy		0.106***	0.245***	0.099***	0.025	0.096***	0.055***	0.224***	0.094***	0.028*	0.134***	0.112***
		(0.035)	(0.025)	(0.036)	(0.022)	(0.018)	(0.015)	(0.023)	(0.031)	(0.016)	(0.025)	(0.030)
Literacy		–0.017	–0.091***	0.056	0.081***	0.047**	0.021	–0.016	0.035	0.034**	0.043	0.061**
		(0.034)	(0.027)	(0.036)	(0.021)	(0.020)	(0.016)	(0.023)	(0.026)	(0.015)	(0.026)	(0.031)
Female		–0.133***	–0.334***	–0.203***	–0.078***	–0.092***	–0.117***	–0.048*	–0.111***	–0.110***	–0.131***	–0.108***
		(0.045)	(0.025)	(0.035)	(0.022)	(0.022)	(0.015)	(0.025)	(0.029)	(0.017)	(0.028)	(0.033)
Pot. experience		0.041***	0.037***	0.024***	0.042***	0.035***	0.030***	0.076***	0.036***	0.024***	0.042***	0.046***
		(0.007)	(0.004)	(0.006)	(0.003)	(0.003)	(0.002)	(0.004)	(0.006)	(0.003)	(0.004)	(0.005)
Pot. experience^2		–0.000**	–0.001***	–0.000	–0.001***	–0.001***	–0.001***	–0.001***	–0.001***	–0.000***	–0.001***	–0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: Least squares regressions (weighted by sampling weights). Dependent variable: log gross hourly wage. All country samples include workers with a college degree. Numeracy and literacy scores are standardized with standard deviation 1 across countries. *Pot. experience* is age − years of schooling − 6. Robust standard errors in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. *Data sources:* PIAAC 2011/12 and 2014/15.