Smarter Teachers Matter: International Evidence from Changes in Teacher Skills and Student Achievement

Eric A. Hanushek[1] Sebastian Reinartz[2] Andreas Schleicher[3]
Simon Wiederhold[1,2,4]

[1] Hoover Institution, Stanford University [2] Halle Institute for Economic Research (IWH)
[3] OECD [4] Martin Luther University Halle-Wittenberg

September 2026

Abstract

While teachers are widely recognized as the most important school input, identifying what makes a teacher effective has proved difficult. We use two rounds of the OECD's Survey of Adult Skills (PIAAC) from roughly a decade apart to provide the first international evidence on how the cognitive skills of teachers change over time and whether student achievement changes with them. By linking teacher skills in 25 countries to the math and reading performance of PISA students, we can assess how a country's changing skill composition of the teacher force affects student performance across subjects. The repeated teacher measures permit removing continuing national differences in subject emphasis. Teacher skills move substantially within individual countries, and these movements are not simply reflections of national skill trends. The changes show up in student achievement. Cross-sectional estimates overstate the importance of teacher skills by about half, but the relationship remains significant when identified entirely from changes within countries. We also relate changes in teacher skills to relative teacher pay and to women's opportunities outside teaching.

Keywords: Teacher quality; cognitive skills; student achievement; PIAAC; PISA

JEL Classification: I20, H40, J20

Contact: hanushek@stanford.edu; sebastian.reinartz@iwh-halle.de; andreas.schleicher@oecd.org; simon.wiederhold@iwh-halle.de

I Introduction

It is widely accepted by parents, school officials, and policy makers, and well documented by research, both that teachers are the most important school input and that there is wide variation in effectiveness of teachers even within individual schools (Bacher-Hicks and Koedel, 2023). Moreover, the learning gains from being in a class with a more effective teacher persist into adulthood with increased college attendance and better labor market outcomes (Hanushek and Rivkin, 2012; Chetty, Friedman and Rockoff, 2014). At the same time, identifying the characteristics of an effective teacher is a much more open question. Value-added measures identify effective teachers only after the fact from what their students learn. Differences in effectiveness are not closely related to salaries, credentials, experience beyond the first few years, or educational background (Hanushek and Rivkin, 2006; Jackson, Rockoff and Staiger, 2014). A growing body of research, however, links teachers' subject knowledge and cognitive skills to student learning (Metzler and Woessmann, 2012; Bietenbeck, Piopiunik and Wiederhold, 2018; Hanushek, Piopiunik and Wiederhold, 2019; Bau and Das, 2020). We exploit recently available international data to provide new evidence about the importance of teacher cognitive skills along with the labor market dynamics that determine skills of individuals entering and remaining in teaching.

If teacher cognitive skills shape student outcomes, then these skills become an important object of policy. As economies evolve, relative opportunities and rewards change, affecting relative pay and other attributes in different occupations. This evolution is especially pertinent to the female-dominated profession of teaching. Many countries have seen significant growth in female opportunities outside of schools, and unless the terms of public sector employment of teachers adjust, such macro changes will alter who selects into teaching. Data from the United States, Australia, and Sweden suggest that relative pay and outside opportunities are directly linked to the skill distribution in teaching (Corcoran, Evans and Schwab, 2004; Hoxby and Leigh, 2004; Bacolod, 2007; Leigh and Ryan, 2008; Grönqvist and Vlachos, 2016). But this evidence comes from individual countries without confirmation that teacher skills matter for students and without any way to easy generalization.

Understanding both the role of teacher skills and the dynamics of the teacher skill distribution is difficult. Within an individual country, both the range of teacher skills and the institutions of the teacher labor market generally have limited variation. These factors show much more variation across countries, making cross-country comparisons potentially useful. But international comparisons face well-known difficulties in extracting causal impacts of country differences on educational outcomes. These difficulties are magnified because existing international evidence rests on single cross-sections, which

cannot separate teacher skills from other stable national features or show how teacher skill patterns change (Golsteyn, Vermeulen and de Wolf, 2016; Hanushek, Piopiunik and Wiederhold, 2019; Estrada and Lombardi, 2023).

Data from two rounds of the OECD's Survey of Adult Skills (PIAAC) provide comparable national assessments of teacher numeracy and literacy roughly a decade apart in 25 countries. We can link the teacher data to the mathematics and reading achievement of about 840,000 students in four waves of the Programme for International Student Assessment (PISA). This opens up two previously unanswered questions. How much do the cognitive skills of a nation's teachers change over time? And if they change, does student achievement change with them?

We address the difficulties of international comparison by linking student performance to country-specific information about teacher cognitive skills and by adapting a within-student, across-subject design to the international setting. By comparing each student's performance across subjects (math and reading here), student-specific factors and other inputs that affect subjects symmetrically are removed (Dee, 2005, 2007; Clotfelter, Ladd and Vigdor, 2010; Lavy, 2015; Bietenbeck, Piopiunik and Wiederhold, 2018; Bietenbeck, Irmert and Sepahvand, 2023). With comparable information about teacher numeracy and literacy across countries, we can estimate the importance of teacher cognitive skills. This design cannot, however, remove continuing subject differences within a country. A nation with a long tradition of emphasizing mathematics may have both more numerate teachers and stronger mathematics students without one causing the other. Observing teacher skills at two points in time lets us test this directly, by asking whether countries whose teachers become relatively more numerate also see their students become relatively stronger in mathematics. It also lets us investigate how economic forces alter the skills of teachers.

The cognitive skill composition of teaching is not fixed, and selection into teaching can change materially within a decade. While average teacher skills change little across the two PIAAC cycles, there are large movements within individual countries. Some countries experience declines on the order of one cross-country standard deviation and others record sizable gains. Moreover, these changes are not simply reflections of population-wide skill trends. Adult skills also evolve, but changes in parent skills and in teacher skills are only moderately correlated. For example, Japanese teachers improve over time in both numeracy and literacy relative to adults in the population, while the opposite is true in Finland and Sweden.

These changes in the teacher force show up in student achievement. In the basic within-student, across-subject estimates, a one-unit shift in the teacher skill profile toward numeracy—where numeracy and literacy are each standardized before differencing—is

associated with a 0.166-unit increase in student performance in math relative to reading. again based on separately standardized scores. [In natural units: a shift of X PIAAC points toward numeracy corresponds to Y PISA points in the math–reading skill difference.] This is comparable to prior cross-country estimates (e.g., Hanushek, Piopiunik and Wiederhold, 2019), but it does not allow for any systematic differences in math versus literacy profiles and preferences of nations. The availability of two comparable observations of teacher skills allows us to remove country fixed effects that control for time-invariant influences of economies, cultures, and educational institutions. Adding country fixed effects reduces the estimated teacher impact to 0.077, but the importance of teacher cognitive skills remains statistically significant. No single country drives the within-country estimate. Country-level teacher skills are measured with high reliability, and any remaining sampling noise attenuates rather than inflates the within-country estimate, so we regard it as conservative. While underlying but unobserved country factors clearly show up in cross-sectional achievement differences, the skill profiles of a country's teachers are an important component of student outcomes. At the level of national teaching workforces, cognitive skills are thus a characteristic that can be observed before anyone enters a classroom and that systematically predicts how much students learn. These findings remain after a variety of robustness checks, and placebo checks [to be provided] reinforce the importance of teacher skills as opposed to general country differences.

Finally, we are able to consider what drives the skill-composition of the teacher force. We relate changes in teacher skills to changes in relative teacher pay and in the opportunities available to women in the economy [to be supplied].

We make three distinct contributions. First, we provide the first international evidence on the evolution of directly assessed teacher cognitive skills. Here we show that the skill composition of the teaching profession can change markedly within a decade. Second, we show that the relationship between teacher and student skills survives a considerably more exacting test than the cross-sectional comparisons on which international evidence has relied. By estimating the existing approaches alongside ours, we also show that cross-sectional estimates overstate the relationship by about half. Third, we contribute to the economics of teacher labor market dynamics and of occupational selection. Repeated PIAAC data allow us to move from cross-sectional comparisons to within-country changes in relative teacher pay, women's access to high-skill occupations outside teaching, and other labor-market conditions.

Section II reviews related literatures. Section III describes PIAAC and PISA and documents the evolution of teacher, adult, parent, and student skills. Section IV develops the empirical framework. Section V presents the student-achievement results and robustness analyses. Section VI studies young-adult skill formation. Section VII examines the

determinants of changes in teacher skills. Section VIII concludes.

II Related Literature

A Teacher Effectiveness and Observable Teacher Quality

A large literature documents substantial variation in teacher effectiveness. Value-added estimates show economically meaningful differences in teachers' contributions to student test scores, while quasi-experimental evidence links exposure to higher-value-added teachers to longer-run educational and labor-market outcomes (Chetty, Friedman and Rockoff, 2014). Teacher effects extend beyond contemporaneous test scores: teachers also affect attendance, behavior, course performance, and educational attainment, and these dimensions of effectiveness need not be strongly correlated with test-score value added (Jackson, 2018). Recent reviews emphasize both the broad evidence that teachers matter and the methodological progress in estimating teacher value added across settings (Bacher-Hicks and Koedel, 2023). Evidence from outside the United States likewise finds substantial teacher effects (Araujo et al., 2016; Bau and Das, 2020), and recent work shows meaningful heterogeneity in teacher value added across student groups (Barrios-Fernández and Riudavets-Barcons, 2024).

A persistent challenge is to identify observable teacher attributes that explain these effects. Formal qualifications, degrees, and many conventional credentials generally account for limited variation in effectiveness (Hanushek and Rivkin, 2006; Jackson, Rockoff and Staiger, 2014). More direct measures of knowledge and cognitive ability are more promising. In Pakistan, teacher content knowledge predicts value added even though observed teacher characteristics overall explain less than 5 percent of its variation (Bau and Das, 2020). In Peru, logical-reasoning and curricular/pedagogical-knowledge assessments are among the teacher selection instruments most strongly associated with subsequent value added (Bertoni et al., 2024). Classroom-process measures also predict student learning in settings with quasi-random assignment of students to teachers (Araujo et al., 2016). Our analysis differs from this work in unit of observation and objective: rather than explaining variation in individual teacher value added, we study whether the measured cognitive skill composition of national teaching workforces changes over time and whether those changes map into changes in national student achievement.

B Changes in Teacher Skills over Time

A related stream of research shows that the skill composition of the teaching profession can change substantially over time. Most of this evidence comes from individual countries and studies changes in the academic aptitude, general cognitive ability, or selection

of individuals entering teaching. In the United States, Corcoran, Evans and Schwab (2004), Hoxby and Leigh (2004), and Bacolod (2007) document pronounced changes in the academic quality of teachers as labor-market opportunities and occupational selection evolved. Leigh and Ryan (2008) provides analogous evidence for Australia, while Grönqvist and Vlachos (2016) documents declining cognitive and social abilities and academic achievement among cohorts entering teaching in Sweden. Taken together, this literature establishes an important fact for our analysis: the skills of the teaching workforce are not a fixed characteristic of national education systems. At the same time, these studies focus primarily on broad measures of aptitude or cognitive ability within individual countries rather than on comparable, domain-specific teacher skills measured repeatedly across countries.

C Teacher Cognitive Skills, Subject Knowledge, and Cross-Subject Designs

Evidence on teachers' cognitive skills is more mixed when the skill measure is broad and domain-general. A systematic review by Bardach and Klassen (2020) concludes that general intelligence measures have weak relationships with teacher effectiveness, while proxies for cognitive skills show at most modest positive associations; numerical ability appears more promising. By contrast, studies that measure subject-specific teacher knowledge generally find clearer relationships with student learning. Metzler and Woessmann (2012) for Peru as well as Bietenbeck, Piopiunik and Wiederhold (2018) and Bietenbeck, Irmert and Sepahvand (2023) for African countries find that teacher subject knowledge is positively related to student achievement in the corresponding subject. Cross-country PIAAC evidence similarly shows a positive relationship between teachers' numeracy and literacy skills and student performance in mathematics and reading (Hanushek, Piopiunik and Wiederhold, 2019).

These studies are part of a broader literature exploiting within-student, across-subject comparisons to absorb student fixed heterogeneity and common educational inputs (Dee, 2005, 2007; Clotfelter, Ladd and Vigdor, 2010; Lavy, 2015). The design is powerful because it compares outcomes within the same student, but it does not automatically remove factors that differ by subject. In international applications, persistent country-by-subject factors—for example, a long-standing mathematics-oriented curriculum or a national tradition emphasizing reading—can remain correlated with both relative teacher inputs and relative achievement. This distinction is central to our contribution. With two PIAAC cycles, we can add country fixed effects to the skill-difference model and thereby absorb all persistent country-specific determinants of the mathematics–reading gap. The resulting within-country estimate provides a direct diagnostic for the magnitude

of persistent subject-specific confounding in conventional cross-subject comparisons.

D Teacher Labor Markets and Selection into Teaching

A separate literature studies how labor-market conditions shape who enters and remains in teaching. In the United States, the opening of professional opportunities to women altered the relative attractiveness of teaching and the skill distribution of female teachers (Corcoran, Evans and Schwab, 2004; Bacolod, 2007). Labor-market conditions at career entry can also affect the quality of teachers who enter the profession (Nagler, Piopiunik and West, 2020). Compensation matters for both sorting and retention: evidence from Wisconsin shows that more flexible pay increased salaries for high-quality teachers, altered mobility and effort, and improved student achievement (Biasi, 2021).

International evidence points to similarly important selection margins. Hanushek, Piopiunik and Wiederhold (2019) show that teacher cognitive skills vary substantially across countries and relate these differences to national teacher labor markets. Countries in which highly skilled women have stronger professional opportunities outside teaching tend to have less-skilled teacher workforces, while higher teacher wage premiums are associated with higher teacher cognitive skills; these wage premiums are also positively related to student achievement. Thus, the paper provides cross-country evidence that both outside options and relative compensation are linked to the skill composition of the teaching profession. Related work shows that teacher cognitive skills differ substantially relative to those of other tertiary-educated workers across countries (Estrada and Lombardi, 2023). In Latin America, relative teacher wages decline with worker skills, consistent with selection incentives that vary across the skill distribution (Estrada and Lombardi, 2023). Other cross-country work relates teacher pay and occupational status directly to student achievement (Dolton and Marcenaro-Gutierrez, 2011; Dolton et al., 2018).

The central limitation of this international evidence is its cross-sectional nature. Persistent country characteristics may jointly shape teacher labor markets, teacher skills, and student achievement. Our repeated international data allow us to move from cross-country differences to within-country changes: we ask whether *changes* in relative teacher pay, women's outside options, and other teacher-labor-market conditions map into changes in teachers' cognitive skills and their relative position in the high-skilled workforce.

III International Comparative Data

A Teacher Cognitive Skills in PIAAC

Our measures of teacher cognitive skills come from the OECD Programme for the International Assessment of Adult Competencies (PIAAC). Cycle 1 was fielded in several rounds beginning in 2011–2012, while Cycle 2 was conducted in 2022–2023. PIAAC directly assesses adult numeracy and literacy on psychometrically linked 500-point scales. We focus on 25 countries observed in both cycles and for which the required PISA student data are available. The repeated country structure is essential: it allows us to distinguish persistent cross-country differences in teacher skills from changes within countries over roughly a decade.

We define teachers as respondents whose current four-digit occupation identifies them as primary-school, secondary-school, or other school teachers. University professors and vocational-school teachers are excluded, as are kindergarten teachers in the baseline definition. Because PIAAC does not report the subject taught by each teacher, our measures describe the numeracy and literacy skills of the national teacher workforce rather than subject-specific teacher groups. Country-by-cycle aggregates use PIAAC final sample weights. The baseline measure is the weighted median, which is less sensitive to extreme observations in cells with moderate teacher samples; we report mean-based measures as robustness checks.

For the econometric analysis, teacher numeracy and literacy are standardized separately using pooled country-by-cycle distributions across both PIAAC cycles. We do not standardize within cycle, because doing so would mechanically remove changes in the cross-country location and dispersion of teacher skills. The resulting numeracy and literacy measures therefore remain on a common scale over time. In the skill-difference models we subtract standardized literacy from standardized numeracy; the difference itself is *not* re-standardized.

Measurement and validation. Country-by-cycle teacher samples are necessarily smaller than the full PIAAC samples, raising a natural concern about sampling noise. We address this in three ways. First, earlier validation exercises using much larger national data sets for the United States and Germany place teachers at similar points of the national skill distribution as PIAAC (Hanushek, Piopiunik and Wiederhold, 2019). Second, we apply the empirical-Bayes framework of Walters (2024) to country-by-cycle mean teacher skills. The estimated signal variance exceeds average sampling variance by more than a factor of twenty, implying average reliability of about 0.96 in both numeracy and literacy. Shrinking noisy country estimates toward their cycle-specific prior means

therefore changes the level estimates only modestly.[1] Third, we reweight Cycle 2 teachers to reproduce the Cycle 1 composition by teacher type, age, gender, and age-by-gender. Our most flexible inverse-probability-weighting adjustment changes Cycle 2 country means by only 2.6 points in numeracy and 2.4 points in literacy on average; adjusted and unadjusted means correlate at 0.99 in both domains. Appendix Section A reports these composition-adjustment results.

Levels and changes. The two tables below report teacher cognitive skills by country in PIAAC Cycles 1 and 2. Cross-country averages move little: median numeracy declines from about 294 to 293 and median literacy from about 297 to 295. This aggregate stability conceals substantial country heterogeneity. Some countries gain strongly in one or both domains, while others experience declines that are large relative to the cross-country distribution. Figure 1 displays the country-specific changes and, importantly for the empirical design below, shows that changes in numeracy and literacy are related but far from mechanically identical. Teacher skill profiles therefore shift both in their overall level and in their relative numeracy–literacy orientation.

[1]The corresponding empirical-Bayes exercise for the numeracy–literacy skill difference will be added once the required restricted-use sampling variances are available for all country-by-cycle cells.

Table 1: Teacher Cognitive Skills and Relative Position, PIAAC Cycle 1

	Pooled	AUT	BEL	CAN	CHL	CZE	DEU	DNK	ESP
Teacher cognitive skills (PIAAC points)									
Numeracy	**294**	300	308	292	262	305	308	295	283
Literacy	**297**	292	303	307	263	300	301	288	290
Relative skill profile									
Numeracy − literacy	−3	8	5	−15	−1	5	7	7	−7
Position in adult skill distribution (percentile)									
Numeracy	**68**	69	68	67	81	73	72	56	75
Literacy	**70**	70	71	72	79	77	74	60	80
Sample size									
Teachers (N)	**5,419**	188	215	834	106	141	127	413	183

	EST	FIN	FRA	GBR	IRL	ISR	ITA	JPN	KOR
Teacher cognitive skills (PIAAC points)									
Numeracy	285	317	302	289	295	270	273	311	287
Literacy	294	322	296	299	300	281	279	319	296
Relative skill profile									
Numeracy − literacy	−9	−5	6	−10	−4	−12	−5	−8	−9
Position in adult skill distribution (percentile)									
Numeracy	60	73	80	65	75	57	67	70	72
Literacy	69	74	77	67	74	62	73	67	74
Sample size									
Teachers (N)	188	221	163	310	180	250	124	147	217

	LTU	NLD	NOR	NZL	SGP	SVK	SWE	USA
Teacher cognitive skills (PIAAC points)								
Numeracy	285	304	302	297	303	294	306	284
Literacy	282	308	304	310	300	290	307	301
Relative skill profile								
Numeracy − literacy	3	−4	−2	−12	3	4	−1	−17
Position in adult skill distribution (percentile)								
Numeracy	66	63	65	64	72	66	62	70
Literacy	64	67	68	71	76	60	65	71
Sample size								
Teachers (N)	133	197	279	198	193	133	147	132

Notes: Teacher cognitive skills are weighted country-specific medians for primary-school, secondary-school, and other teaching professionals. The relative skill profile is median numeracy minus median literacy. Percentile ranks report the position of the median teacher in the corresponding skill distribution of adults aged 25–65, excluding teachers. In Finland, occupations are available only at the two-digit level, so the teacher sample comprises all teaching professionals (ISCO-08 code 23). Observations report the unweighted number of teachers used to construct the country-level measures. Individuals are weighted using PIAAC final sample weights. *Data source:* PIAAC Cycle 1 (2012, 2015).

Table 2: Teacher Cognitive Skills and Relative Position, PIAAC Cycle 2

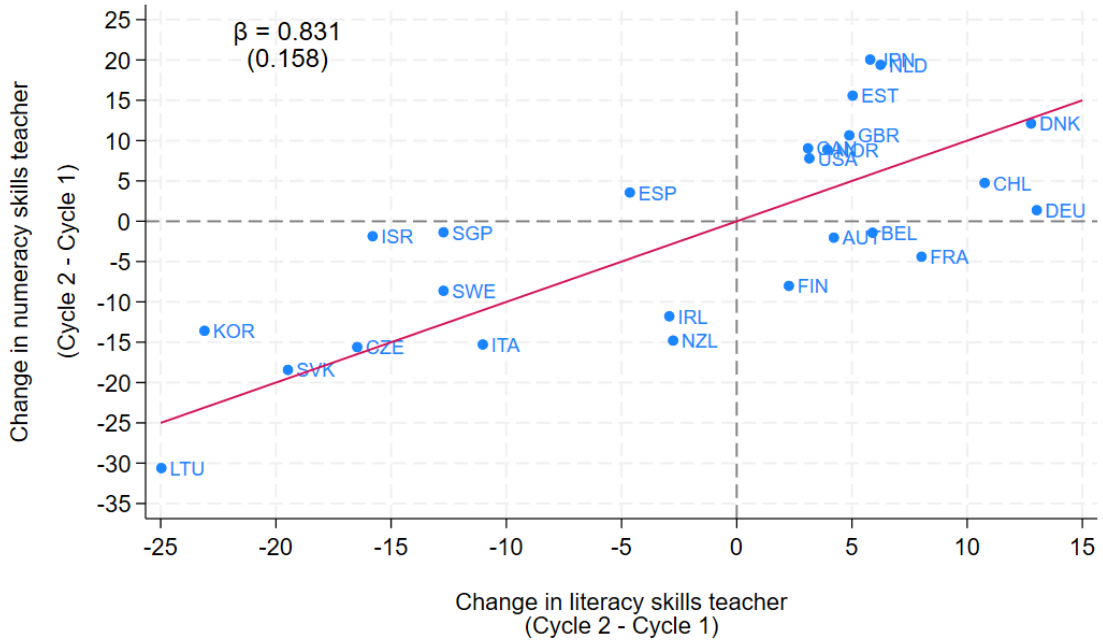
	Pooled	**AUT**	**BEL**	**CAN**	**CHL**	**CZE**	**DEU**	**DNK**	**ESP**
Teacher cognitive skills (PIAAC points)									
Numeracy	**293**	298	307	301	267	289	310	307	286
Literacy	**295**	297	309	310	274	283	314	301	285
Relative skill profile									
Numeracy − literacy	−2	2	−2	−9	−7	6	−5	6	1
Position in adult skill distribution (percentile)									
Numeracy	**67**	68	64	68	83	66	72	62	75
Literacy	**69**	75	66	74	82	63	77	64	77
Sample size									
Teachers (N)	**4,511**	180	194	301	133	181	140	222	224

	EST	**FIN**	**FRA**	**GBR**	**IRL**	**ISR**	**ITA**	**JPN**	**KOR**
Teacher cognitive skills (PIAAC points)									
Numeracy	300	309	297	300	284	268	258	331	273
Literacy	299	324	304	304	297	266	268	325	273
Relative skill profile									
Numeracy − literacy	2	−15	−7	−4	−13	2	−9	7	0
Position in adult skill distribution (percentile)									
Numeracy	59	56	76	69	70	62	62	77	69
Literacy	63	61	81	74	74	62	66	74	70
Sample size									
Teachers (N)	221	179	174	166	168	365	107	170	211

	LTU	**NLD**	**NOR**	**NZL**	**SGP**	**SVK**	**SWE**	**USA**
Teacher cognitive skills (PIAAC points)								
Numeracy	255	323	311	283	302	276	297	292
Literacy	257	314	308	307	287	271	295	304
Relative skill profile								
Numeracy − literacy	−2	9	3	−24	15	5	3	−12
Position in adult skill distribution (percentile)								
Numeracy	63	71	61	66	62	62	53	70
Literacy	66	70	63	74	66	60	51	72
Sample size								
Teachers (N)	108	106	177	132	144	245	133	130

Notes: Teacher cognitive skills are weighted country-specific medians for primary-school, secondary-school, and other teaching professionals. The relative skill profile is median numeracy minus median literacy. Percentile ranks report the position of the median teacher in the corresponding skill distribution of adults aged 25–65, excluding teachers. Observations report the unweighted number of teachers used to construct the country-level measures. Individuals are weighted using PIAAC final sample weights. *Data source:* PIAAC Cycle 2 (2022, 2023).

Figure 1: Changes in Teacher Cognitive Skills



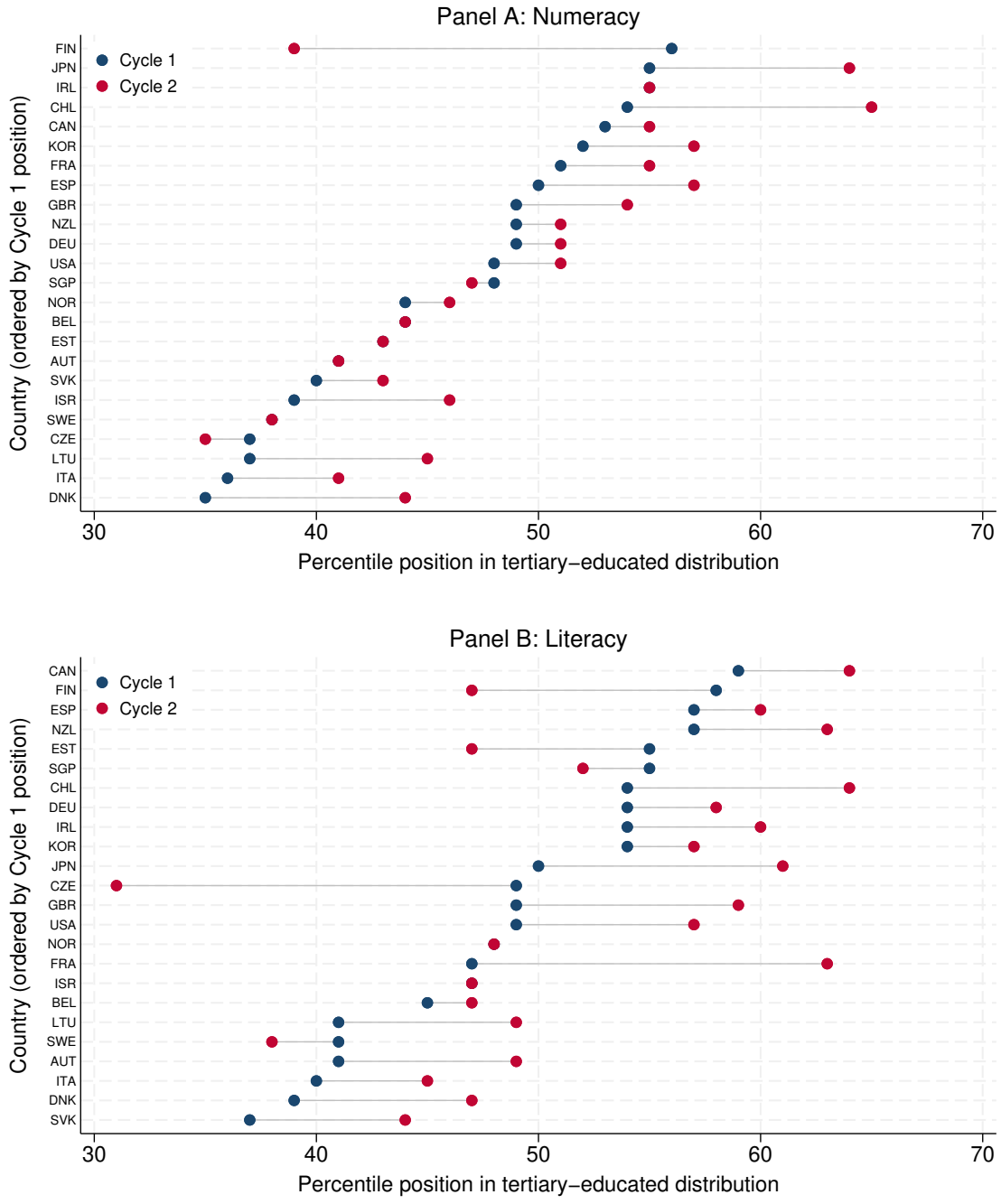
Notes: The figure shows country-level changes in median teacher numeracy and literacy between PIAAC Cycles 1 and 2. The 45-degree line denotes equal changes in both domains; deviations from that line indicate shifts in the relative teacher skill profile. *Data sources:* PIAAC Cycles 1 and 2.

Selection into teaching. Absolute teacher skills combine the overall skill level of a country with occupational selection into teaching. To isolate the latter dimension descriptively, Figure 2 places teachers in the skill distribution of tertiary-educated nonteachers in the same country and cycle. The movements are large. Across countries, teachers' percentile positions change by as much as 17 points downward and 11 points upward in numeracy, and by 18 points downward and 16 points upward in literacy. Finland's numeracy position, for example, falls by 17 percentile points, whereas Chile's rises by 11; in literacy, the Czech Republic falls by 18 points and France rises by 16. These shifts show that changes in measured teacher skills are not simply movements of the entire high-skilled population: the relative selection of workers into teaching changes materially as well.

B Adult and Parent Skill Environments

Teacher-skill changes may reflect broader changes in national skill distributions rather than teacher-specific developments. We therefore construct two comparison measures. First, adult skills are the median numeracy and literacy proficiency of adults aged 25–65. Across the 25-country sample, adult numeracy changes little on average, while adult literacy declines more noticeably. The resulting teacher–adult skill differences also move substantially across countries (Appendix Figures B1–B4). These patterns motivate a

Figure 2: Teacher Percentile Position in the Tertiary-Educated Skill Distribution



Notes: The figure shows teachers' percentile position within the skill distribution of tertiary-educated adults excluding teachers. Panel A reports numeracy and Panel B literacy. Lines connect each country's Cycle 1 and Cycle 2 positions; countries are ordered by their Cycle 1 position. *Data sources:* PIAAC Cycles 1 and 2.

robustness specification that controls directly for the contemporaneous numeracy–literacy skill profile of nonteacher adults.

Second, because family inputs may affect mathematics and reading differently, we construct predicted parental cognitive skills for each PISA student. We use PIAAC respondents aged 35–59 with children and estimate country-specific prediction models for numeracy and literacy based on gender, educational attainment, and books in the household—characteristics observed for parents in PISA as well. We apply these models to the corresponding characteristics of students' mothers and fathers and use the higher predicted parental skill as our baseline family-skill measure.

Predicted parent numeracy rises by about 1.4 PIAAC points on average between cycles, whereas predicted parent literacy declines by about 5.3 points. The parent numeracy-minus-literacy difference consequently widens. Importantly, teacher and parent skill changes are related but not close to one-for-one: across countries, their changes correlate at 0.43 in numeracy and 0.54 in literacy. This imperfect comovement is useful for identification because it allows us to distinguish changes in the teacher skill profile from contemporaneous changes in family skill inputs.

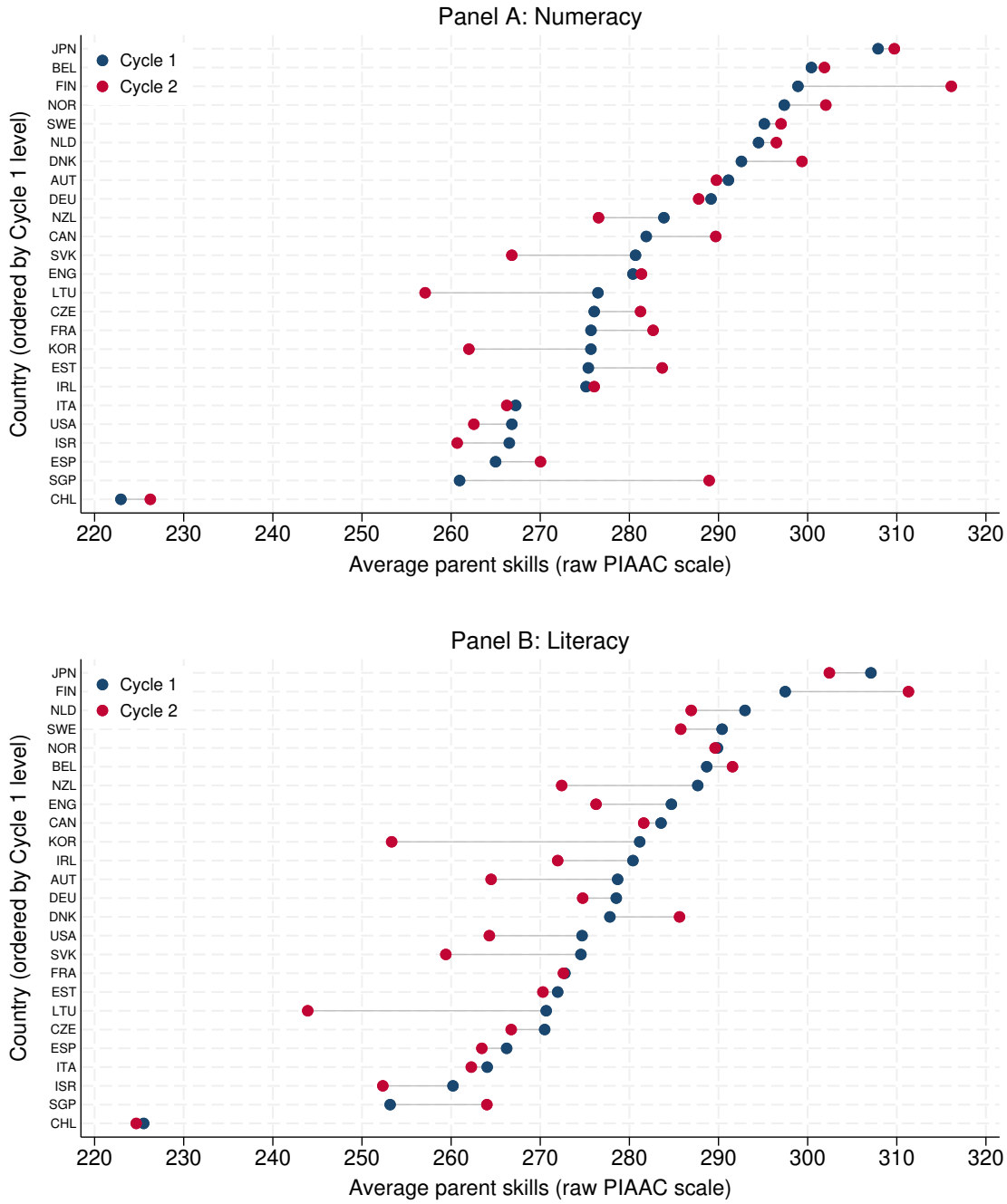
C Student Achievement in PISA

Student achievement comes from the OECD Programme for International Student Assessment (PISA), which assesses 15-year-old students in mathematics and reading. We use PISA 2009 and 2012 for the period matched to PIAAC Cycle 1 and PISA 2018 and 2022 for the period matched to Cycle 2, yielding roughly 840,000 student observations across the 25 countries. Using two PISA waves per PIAAC cycle improves precision and reduces dependence on any single PISA cohort, although the temporal match between teachers and students is necessarily imperfect. We therefore report alternative PISA–PIAAC alignments as robustness checks.

PISA provides rich student, family, and school information. Student controls include age, gender, migrant generation, and language spoken at home; family controls include books, parental education, and parental occupation; school controls include location, private-school status, and school size. We additionally use domain-specific instructional time where it is comparably measured. The PISA 2022 questionnaire does not provide a language-instruction measure comparable to earlier waves; the current data construction carries forward the 2018 country value for that component. We therefore do not rely on this imputed measure for identification: specifications without instructional-time controls and specifications excluding PISA 2022 are important robustness checks.

Figure 4 documents broad declines in student achievement across the two periods, partic-

Figure 3: Predicted Parent Cognitive Skills Across PIAAC Cycles



Notes: The figure shows country-level predicted parent skills in the two PIAAC cycles. Panel A displays numeracy and Panel B literacy. Country values are weighted means across PISA students. *Data sources:* PIAAC Cycles 1 and 2 and PISA 2009, 2012, 2018, and 2022.

ularly in mathematics. The empirical analysis does not, however, rely on these common trends. Mathematics and reading scores are standardized separately on pooled, time-invariant scales across the included PISA waves, and the skill-difference outcome is then defined as standardized mathematics minus standardized reading. The difference itself is *not* re-standardized. Hence, our preferred model asks whether countries with a relative shift in teacher skills toward numeracy also experience a relative shift in student performance toward mathematics.

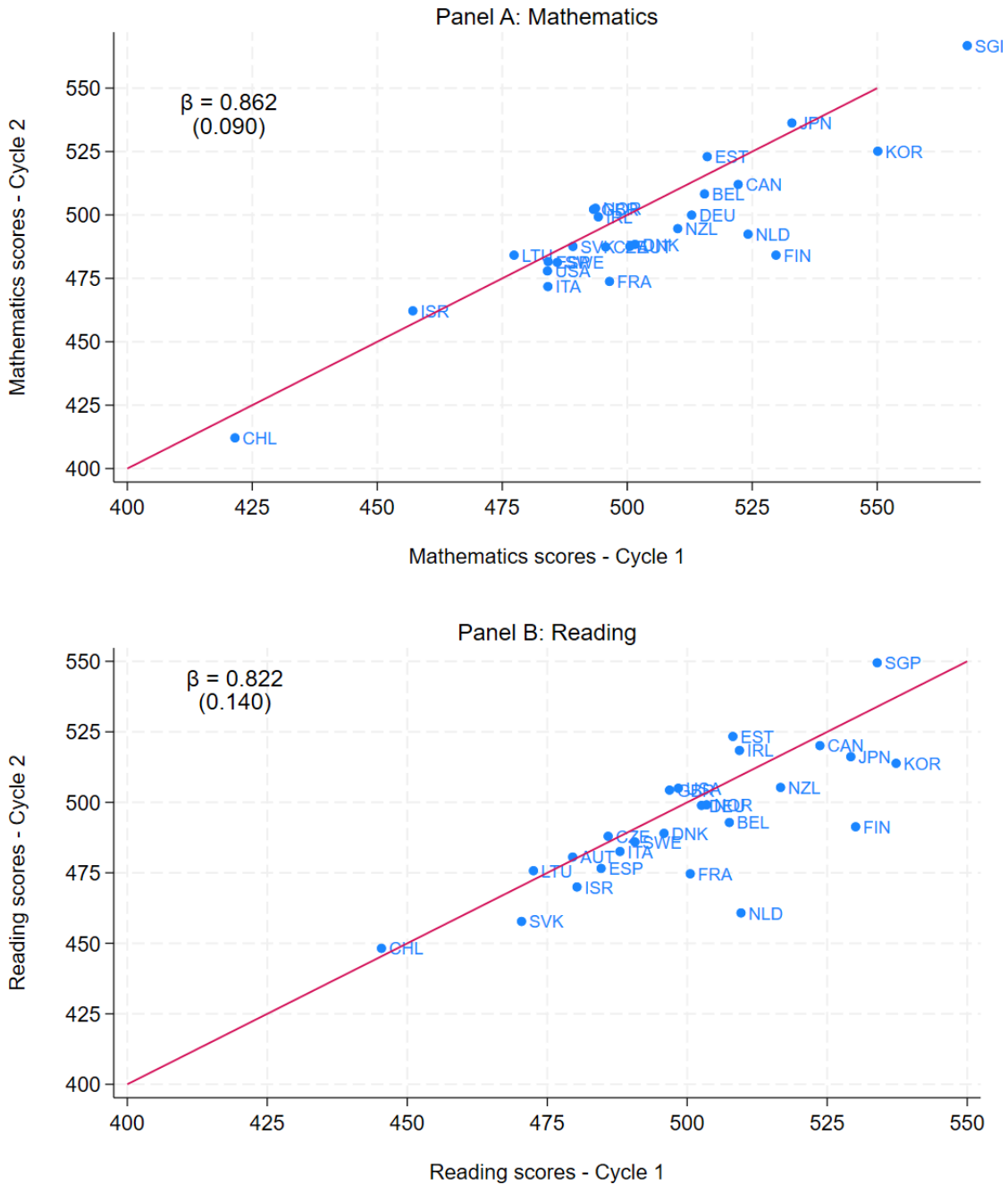
IV Conceptual Framework and Empirical Strategy

We organize the empirical analysis around three nested specifications. The levels model relates teacher numeracy to student mathematics and teacher literacy to student reading. The standard skill-difference model then compares mathematics with reading within the same student and numeracy with literacy within the same teacher workforce, removing factors that affect the two domains symmetrically. Finally, the repeated PIAAC structure allows us to add country fixed effects to this cross-subject comparison. This third specification is the paper's preferred design: it removes persistent country-by-subject heterogeneity and identifies the teacher-skill coefficient from within-country changes in relative teacher and student skill profiles. The progression from the second to the third model is therefore not merely a robustness exercise; it directly measures how much of the conventional cross-subject relationship is attributable to persistent subject-specific country factors.

Let A^k_{icw} denote achievement of student i in country c, wave $w \in \{2009, 2012, 2018, 2022\}$, and student achievement domain $k \in \{M, R\}$, where M denotes mathematics and R denotes reading. Let $T^N_{c,s(w)}$ and $T^L_{c,s(w)}$ denote the cognitive skills of the teacher workforce in the numeracy and literacy domains, respectively. The index $s(w) \in \{1, 2\}$ denotes the PIAAC cycle matched to PISA wave w, with $s(2009) = s(2012) = 1$ and $s(2018) = s(2022) = 2$. We refer to mathematics–numeracy and reading–literacy as the aligned domain pairs. Because PIAAC does not allow us to link individual teachers to individual students, $T^N_{c,s(w)}$ and $T^L_{c,s(w)}$ are measured at the country-by-cycle level and should be interpreted as characteristics of the national teacher pool.[2] A flexible two-domain education production framework is

[2]Two PISA waves are matched to each PIAAC cycle, so teacher skills are constant across those two PISA waves within a country.

Figure 4: Student Performance Across the Two PIAAC-Matched Periods



Notes: The figure compares country-level student performance in mathematics (Panel A) and reading (Panel B) across the two PIAAC-matched periods. The 45-degree line denotes no change. *Data sources:* PISA 2009, 2012, 2018, and 2022.

$$A^M_{icw} = \alpha^M + \beta_{MN} T^N_{c,s(w)} + \beta_{ML} T^L_{c,s(w)} + X^{M\prime}_{icw}\gamma^M + \mu_{icw} + \kappa_{cw} + \eta^M_c + \lambda^M_w + \zeta^M_{cw} + \varepsilon^M_{icw}, \tag{1}$$

$$A^R_{icw} = \alpha^R + \beta_{RN} T^N_{c,s(w)} + \beta_{RL} T^L_{c,s(w)} + X^{R\prime}_{icw}\gamma^R + \mu_{icw} + \kappa_{cw} + \eta^R_c + \lambda^R_w + \zeta^R_{cw} + \varepsilon^R_{icw}. \tag{2}$$

The coefficients β_{MN} and β_{RL} are aligned-domain effects: teacher numeracy on student mathematics and teacher literacy on student reading. The coefficients β_{ML} and β_{RN} allow for cross-domain spillovers. For example, teacher literacy may matter for mathematics because students must understand word problems, while teacher numeracy may matter for reading if analytical reasoning supports text comprehension. The vectors X^M_{icw} and X^R_{icw} contain observable control variables, with coefficient vectors γ^M and γ^R that are allowed to differ across domains. The terms μ_{icw} and κ_{cw} capture factors that affect both achievement domains symmetrically. At the student level, μ_{icw} includes general ability, motivation, test-taking effort, and family background. At the country-by-wave level, κ_{cw} captures broad educational investments, general school quality, or macroeconomic shocks that affect mathematics and reading similarly. The terms η^k_c denote time-invariant country-by-domain factors, such as persistent differences in curricula, tracking regimes, or national traditions that favor one domain over the other. The terms λ^k_w denote common wave-by-domain shocks, such as changes in the relative scaling or difficulty of mathematics and reading tests. Finally, ζ^k_{cw} captures time-varying country-by-domain shocks, the residual country-by-domain component that remains after η^k_c and λ^k_w, such as mathematics reforms, reading campaigns, changes in domain-specific instructional time, or other domain-specific policy changes.

A Baseline Levels Model

The most direct empirical model relates achievement in each domain to teacher skills in the aligned domain and exploits the full student-level variation:

$$A^k_{icw} = \theta^k T^{m(k)}_{c,s(w)} + X^{k\prime}_{icw}\gamma^k + \lambda^k_w + e^k_{icw}, \tag{3}$$

where $m(M) = N$ and $m(R) = L$. Thus, student mathematics is related to teacher numeracy, while student reading is related to teacher literacy. This model asks whether students perform better in countries and cycles in which teachers have higher cognitive skills in the relevant domain. We stack PISA student microdata from the waves closest to PIAAC Cycle 1 (PISA 2009 and 2012) and Cycle 2 (PISA 2018 and 2022), and merge

country×cycle×domain teacher skill measures to each student observation. We estimate the model separately for each domain and include an extensive set of control variables X_{icw}^k at multiple levels. Student characteristics include age, gender, migrant status (first-generation or second-generation), and language spoken at home. Parent characteristics comprise parents' educational degree, number of books at home, and occupation. School characteristics include school location, number of students per school, as well as private school status. Country characteristics include expenditures per student, school starting age and weekly instructional time in math and language classes. All regressions include indicators for imputation flags and PISA-wave fixed effects λ_w^k that absorb global time trends common to all countries. We cluster standard errors at the country level because the key regressor varies only at the country×cycle×domain level and each country enters with both cycles, and we use wild cluster bootstrap as a robustness check of the results.

The levels model is informative, but it requires strong identifying assumptions. Even with rich controls, θ^k may capture unobserved country-level factors that are correlated with both teacher skills and student achievement. Let B_{cw}^k denote omitted determinants of student achievement in domain k, such as educational culture, institutional quality, curriculum rigor, long-run school-system features, or time-varying country shocks. Then the probability limit of the levels estimator can be written as

$$\operatorname{plim} \widehat{\theta}^k_{\text{levels}} = \theta^k + \frac{\operatorname{Cov}\left(T^{m(k)}_{c,s(w)}, B^k_{cw}\right)}{\operatorname{Var}\left(T^{m(k)}_{c,s(w)}\right)}, \tag{4}$$

where $T^{m(k)}_{c,s(w)}$ is understood as residualized with respect to included controls and wave fixed effects. The levels model may therefore be biased if countries with more skilled teachers also differ in other achievement-enhancing ways. This concern motivates designs that exploit variation across domains within the same student.

B The Standard Within-Student Skill-Difference Model

The standard within-student, between-subject model exploits the fact that both students and teachers are observed in two domains. Subtracting reading from mathematics in equations (1) and (2) yields

$$
\begin{aligned}
D^{A}_{icw} &\equiv A^{M}_{icw} - A^{R}_{icw} \\
&= (\alpha^{M} - \alpha^{R}) + (\beta_{MN} - \beta_{RN}) T^{N}_{c,s(w)} - (\beta_{RL} - \beta_{ML}) T^{L}_{c,s(w)} \\
&\quad + X^{D\prime}_{icw}\gamma + \eta^{D}_{c} + \lambda^{D}_{w} + \zeta^{D}_{cw} + u^{D}_{icw},
\end{aligned} \tag{5}
$$

where

$$
X^{D\prime}_{icw}\gamma = X^{M\prime}_{icw}\gamma^{M} - X^{R\prime}_{icw}\gamma^{R}, \quad \eta^{D}_{c} = \eta^{M}_{c} - \eta^{R}_{c}, \quad \lambda^{D}_{w} = \lambda^{M}_{w} - \lambda^{R}_{w},
$$

$$
\zeta^{D}_{cw} = \zeta^{M}_{cw} - \zeta^{R}_{cw}, \quad u^{D}_{icw} = \varepsilon^{M}_{icw} - \varepsilon^{R}_{icw}.
$$

Here, $X^{D\prime}_{icw}\gamma$ collects observed characteristics whose effects are allowed to differ across mathematics and reading. In an equivalent stacked specification, this amounts to interacting the covariates with the subject indicator; in the directly differenced specification, the coefficient on each covariate estimates the difference between its mathematics and reading effects. By contrast, the common components μ_{icw} and κ_{cw} drop out by construction. The skill-difference model therefore removes general student ability, broad test-taking factors, and country-by-wave shocks only to the extent that they affect the two domains symmetrically; observable characteristics with differential subject effects should remain in X^{D}_{icw}.

Define the teacher skill-difference and average teacher skill as

$$
D^{T}_{c,s(w)} = T^{N}_{c,s(w)} - T^{L}_{c,s(w)}, \tag{6}
$$

$$
\bar{T}_{c,s(w)} = \frac{T^{N}_{c,s(w)} + T^{L}_{c,s(w)}}{2}. \tag{7}
$$

Let

$$
a = \beta_{MN} - \beta_{RN}, \tag{8}
$$

$$
b = \beta_{RL} - \beta_{ML}. \tag{9}
$$

Here, a is the net effect of teacher numeracy on the student mathematics-minus-reading skill-difference, while b is the net effect of teacher literacy on the student reading-minus-mathematics skill-difference. The teacher-skill component in equation (5) can then be rewritten as

$$aT^{N}_{c,s(w)} - bT^{L}_{c,s(w)} = \frac{a+b}{2} D^{T}_{c,s(w)} + (a-b)\bar{T}_{c,s(w)}. \tag{10}$$

Equation (10) shows the restriction underlying the parsimonious skill-difference model. If

$$\beta_{MN} - \beta_{RN} = \beta_{RL} - \beta_{ML}, \tag{11}$$

then $a = b$, the average teacher skill term drops out, and the student achievement difference depends only on the teacher numeracy-minus-literacy skill-difference. Under this symmetry condition, the standard skill-difference model is

$$D^{A}_{icw} = \beta D^{T}_{c,s(w)} + X^{D\prime}_{icw}\gamma + \lambda^{D}_{w} + v^{D}_{icw}. \tag{12}$$

Here, $v^{D}_{icw} = u^{D}_{icw} + \eta^{D}_{c} + \zeta^{D}_{cw}$, and the constant $(\alpha^{M} - \alpha^{R})$ is absorbed into the intercept. Under this symmetry condition, $\beta = a = b$, which coincides with the general expression for β in equation (16). This model is computationally equivalent to pooling mathematics and reading observations and including student fixed effects. It is therefore closely related to the between-subject designs used in prior work on teachers, instruction, and domain-specific skills, including Dee (2005, 2007), Clotfelter, Ladd and Vigdor (2010), Lavy (2015), Bietenbeck, Piopiunik and Wiederhold (2018), Hanushek, Piopiunik and Wiederhold (2019), and Bietenbeck, Irmert and Sepahvand (2023).

For the skill-difference coefficient to be interpreted as the effect of teacher cognitive skills or the teacher's knowledge in the domain on student cognitive skills, rather than merely as an association between relative skill profiles, additional structure is required. A transparent sufficient set of restrictions is

$$\beta_{MN} = \beta_{RL} \equiv \beta^{O}, \qquad \beta_{ML} = \beta_{RN} \equiv \beta^{C}, \tag{13}$$

where the superscript O denotes own, that is, aligned-domain effects, and C denotes cross-domain effects. The first restriction states that aligned-domain effects are symmetric: teacher numeracy affects student mathematics by the same amount as teacher literacy affects student reading. The second restriction states that cross-domain spillovers are

symmetric: teacher literacy affects student mathematics by the same amount as teacher numeracy affects student reading. Under these restrictions,

$$\beta_{MN} - \beta_{RN} = \beta_{RL} - \beta_{ML} = \beta^O - \beta^C. \tag{14}$$

Thus, the parsimonious skill-difference model identifies the common aligned-domain effect net of symmetric cross-domain spillovers. In the special case of no cross-domain spillovers, $\beta^C = 0$, the skill-difference coefficient identifies the common aligned-domain effect β^O of teacher cognitive skills on student cognitive skills.[3]

If condition (11) does not hold, average teacher skill belongs in the skill-difference equation. The more general skill-difference model is

$$D^A_{icw} = \beta D^T_{c,s(w)} + \rho \bar{T}_{c,s(w)} + X^{D\prime}_{icw}\gamma + \lambda^D_w + v^D_{icw}, \tag{15}$$

where

$$\beta = \frac{(\beta_{MN} - \beta_{RN}) + (\beta_{RL} - \beta_{ML})}{2}, \tag{16}$$

$$\rho = (\beta_{MN} - \beta_{RN}) - (\beta_{RL} - \beta_{ML}). \tag{17}$$

The parsimonious model imposes $\rho = 0$. This motivates specifications that include average teacher skill in addition to the teacher skill-difference. Stability of β across these specifications supports the interpretation that the teacher skill-difference captures the relevant relative skill component rather than asymmetric level effects.

The standard skill-difference model improves on the levels model but does not eliminate all confounding. It removes non-domain-specific student and country factors, but it may still be biased by domain-specific country factors. These include persistent math-oriented curricula, long-standing reading traditions, tracking systems that favor one domain, or cultural preferences for mathematics versus reading. Formally, the probability limit of the standard skill-difference estimator can be expressed as

[3]The skill-difference model does not identify the effect of raising teachers' overall cognitive skill level equally in both domains, because such a common teacher-skill component is differenced out. Instead, it identifies whether variation in domain-specific teacher cognitive skills maps into corresponding variation in domain-specific student achievement.

$$\operatorname{plim} \widehat{\beta}_{\text{diff}} = \beta + \frac{\operatorname{Cov}\left(D^T_{c,s(w)}, u^D_{icw} + \eta^D_c + \zeta^D_{cw}\right)}{\operatorname{Var}\left(D^T_{c,s(w)}\right)}, \tag{18}$$

where $D^T_{c,s(w)}$ is understood as residualized with respect to included controls and wave fixed effects. Equation (18) highlights the central limitation of the standard within-student, between-subject design: countries with a larger teacher numeracy advantage may also have persistent or time-varying domain-specific features that independently raise student mathematics performance relative to reading.

This framework is closely related to the intergenerational analysis of the comparative-skill-advantage framework in Hanushek et al. (2023).[4] The same logic applies here: differencing removes broad factors that affect both domains similarly, but the parsimonious one-difference model is exact only under a symmetry condition. If the symmetry condition fails, average skill terms and the unrestricted two-domain system are required to interpret the skill-difference coefficient.

C Country Fixed Effects Within-Student Skill-Difference Model

The second PIAAC cycle allows us to test a restriction that is typically maintained, rather than examined, in international cross-subject designs. A within-student comparison removes student heterogeneity and common country inputs, but it cannot remove a persistent country-specific orientation toward one subject. With repeated teacher-skill measures, country fixed effects absorb precisely these stable mathematics-versus-reading differences. We estimate

$$D^A_{icw} = \beta D^T_{c,s(w)} + \rho \bar{T}_{c,s(w)} + X^{D\prime}_{icw}\gamma + \lambda^D_w + \delta_c + v^D_{icw}. \tag{19}$$

Country fixed effects absorb all time-invariant determinants of the mathematics-reading skill-difference within a country, including persistent domain-specific curricula, tracking systems, and cultural preferences for mathematics or reading. Wave fixed effects absorb changes in the mathematics-reading skill-difference that are common across countries. With two cycles, equation (19) is equivalent to a difference-in-difference-in-difference

[4] In that setting, individual comparative skill advantage is defined as the difference between math and language ranks. In both settings, the empirical object is a within-unit difference across domains: in the intergenerational setting, the unit is the individual or parent-child pair; in the teacher-skills setting, it is the country-by-wave teacher workforce and the corresponding student achievement skill-difference. In both cases, differencing removes broad factors that affect both domains symmetrically, while the parsimonious one-difference model is exact only under a symmetry condition on the underlying two-domain production framework.

specification, taking differences across domains, over time, and across countries:

$$\Delta D_c^A = \beta \Delta D_c^T + \rho \Delta \bar{T}_c + \Delta X_c^{D\prime} \gamma + \Delta \zeta_c^D + \Delta u_c^D, \tag{20}$$

where $\Delta(\cdot)$ denotes the change from PIAAC Cycle 1 to Cycle 2. The preferred specification therefore asks whether countries in which teachers became relatively more numerate than literate also experienced relative improvements in student mathematics performance compared with reading.

Equation (19) provides a direct diagnostic for persistent subject-specific confounding in the standard cross-subject estimator. If countries with a persistent numeracy advantage among teachers also have persistently mathematics-oriented curricula, institutions, or preferences, the pooled skill-difference coefficient will partly reflect those factors and should attenuate when country fixed effects are added. A positive coefficient that survives this transformation is stronger evidence that changes in the relative skill composition of teachers co-move with changes in the corresponding student achievement profile. Importantly, attenuation itself is informative: it quantifies how much the standard cross-subject comparison loads on persistent country-by-subject heterogeneity.

The country-fixed-effects skill-difference model removes persistent domain-specific confounds, but it does not remove time-varying domain-specific shocks. The remaining identifying assumption is

$$\operatorname{Cov}\left(\Delta D_c^T, \Delta \zeta_c^D \mid \Delta X_c^D, \Delta \bar{T}_c\right) = 0. \tag{21}$$

In words, the remaining threat is a time-varying, domain-specific country shock that jointly changes the teacher skill profile and the student achievement profile. Examples include mathematics-specific reforms, reading campaigns, or changes in subject-specific instructional inputs. We address observable components through parental-skill and instructional-time differences and, in the completed analysis, through a broader set of differential student, family, and school controls. We further use changes in the skill profiles of nonteacher adults and other occupations as falsification and confounding checks.

The stronger specification also places greater demands on measurement. Teacher skills are country-by-cycle aggregates estimated from finite PIAAC samples, and the identifying regressor is a difference across both domains and time. Under classical sampling error, country fixed effects and long differences therefore tend to attenuate the coefficient more strongly than pooled cross-sectional comparisons. This does not remove the need for careful inference or robustness checks, but it implies that a smaller within-country coefficient is not, by itself, evidence against a substantive relationship. We address this issue with empirical-Bayes diagnostics, alternative teacher-skill measures, composition adjustments,

and sample-restriction checks.

A useful way to summarize the empirical contribution is a between–within decomposition of equation (19):

$$\begin{aligned} D^A_{icw} =& \beta_W \left(D^T_{c,s(w)} - \bar{D}^T_c \right) + \beta_B \bar{D}^T_c + \rho_W \left(\bar{T}_{c,s(w)} - \tilde{T}_c \right) + \rho_B \tilde{T}_c \\ &+ \left(X^D_{icw} - \bar{X}^D_c \right)' \gamma_W + \bar{X}^{D\prime}_c \gamma_B + \lambda^D_w + v^D_{icw}. \end{aligned} \tag{22}$$

Here, $\bar{D}^T_c$ denotes the country mean of the teacher skill-difference across the two cycles, $\tilde{T}_c$ the country mean of average teacher skill across the two cycles, and $\bar{X}^D_c$ the country mean of the controls. The bar in $\bar{D}^T_c$ and $\bar{X}^D_c$ and the tilde in $\tilde{T}_c$ denote averages over time, whereas the bar in $\bar{T}_{c,s(w)}$ denotes an average over the two teacher skill domains. Country fixed effects are omitted because $\bar{D}^T_c$, $\tilde{T}_c$, and $\bar{X}^D_c$ are country-level and would be collinear with them. The between coefficient β_B captures the association across countries in their average relative teacher-skill profiles, conditional on the corresponding between-country components of the controls. The within coefficient β_W captures the association between changes in relative teacher skills and relative student achievement within countries and, in the fully decomposed specification, corresponds to β in equation (19).

Comparing these coefficients allows us to assess whether the standard within-student, between-subject design is mainly picking up persistent domain-specific country differences, or whether the relationship also holds when such persistent differences are removed.

The three models thus have distinct identifying content. The levels model provides the most direct but least protected association. The standard skill-difference model removes symmetric student and country factors but still mixes persistent between-country differences with within-country changes. The country-fixed-effects skill-difference model isolates the latter. The between–within decomposition reports both components in a single regression and makes transparent what the country fixed effects remove.

V Teacher Cognitive Skills and Student Performance over Time

We present the estimates in the same sequence as the empirical framework. The levels specification establishes the familiar cross-country relationship between teacher skills and student achievement. The standard skill-difference model removes symmetric student and country factors. The country-fixed-effects skill-difference model then asks whether the relationship survives when persistent country-by-subject differences are absorbed and identification comes from within-country changes.

A Levels: Teacher Skills and Achievement

Table 3 reports the levels estimates. Columns (1) and (5) relate student achievement to teacher skills in the aligned domain. A one-standard-deviation increase in country-level teacher numeracy is associated with 0.162 standard deviations higher mathematics achievement, while a one-standard-deviation increase in teacher literacy is associated with 0.113 standard deviations higher reading achievement. Both estimates are statistically significant.

The remaining columns add student, family, school, and country controls, predicted parent skills, and instructional time. The teacher coefficients decline but remain precisely estimated: 0.102 for mathematics and 0.084 for reading in the richest specifications. The similarity of these magnitudes to earlier cross-sectional estimates is useful as a benchmark, but we do not treat the levels regressions as the paper's preferred evidence because they remain exposed to unobserved country heterogeneity.

Table 3: Student Performance and Teacher Cognitive Skills (Levels Model)

	Student Math Performance				Student Reading Performance			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Teacher cognitive skills	0.162***	0.136***	0.102***	0.102***	0.113***	0.081***	0.083***	0.084***
	(0.050)	(0.036)	(0.027)	(0.026)	(0.029)	(0.018)	(0.018)	(0.017)
Parent cognitive skills			0.043*	0.039**			−0.003	−0.002
			(0.024)	(0.017)			(0.015)	(0.014)
Weekly instruct. time (h/week)				−0.015				0.008
				(0.034)				(0.016)
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	839,277	839,277	839,277	839,277	839,277	839,277	839,277	839,277
Countries	25	25	25	25	25	25	25	25
Adj. R^2	0.03	0.26	0.26	0.27	0.01	0.27	0.27	0.27

Notes: Least squares regressions weighted by students' inverse sampling probability, rescaled to give each country×wave cell equal weight in the pooled sample. Dependent variable: student PISA test score in math (Columns 1–4) and in reading (Columns 5–8), standardized using the pooled mean and standard deviation across all countries and waves (pooled z-score). Country-level teacher cognitive skills refer to numeracy in Columns 1–4 and to literacy in Columns 5–8, and are z-standardized across countries and PIAAC cycles. Parent cognitive skills are computed as the maximum of mother's and father's skills in numeracy (Columns 1–4) or literacy (Columns 5–8), 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 and private-school status. Country characteristics include the school starting age and educational expenditure per student. Weekly instruct. time (h/week) is a country×wave level control measuring average weekly instructional time in math classes (Columns 1–4) or language classes (Columns 5–8), aggregated from school questionnaire responses to the country×wave level; the 2022 wave carries forward 2018 country means. All regressions include controls for respective imputation dummies and dummies 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:* PIAAC Cycle 1 (2012, 2015), PIAAC Cycle 2 (2022, 2023) and PISA 2009, 2012, 2018 and 2022.

B Standard Within-Student Skill-Difference Estimates

Table 4: Within-Student Skill-Difference Model

	(1) Teacher skill diff.	(2) Parent skill diff.	(3) Instr.-time diff.	(4) All diff.	(5) Teacher skill diff. + $\bar{T}$	(6) All diff. + $\bar{T}$
Teacher skill difference	0.187***			0.166***	0.186***	0.166***
	(0.032)			(0.034)	(0.030)	(0.031)
Parent skill difference		0.124***		0.043		0.041
		(0.034)		(0.032)		(0.028)
Instructional-time difference			0.020	0.030		0.022
			(0.065)	(0.047)		(0.044)
$\bar{T}$					0.049*	0.047*
					(0.026)	(0.028)
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Overall R^2	0.03	0.01	0.00	0.03	0.03	0.03
Observations	839,277	839,277	839,277	839,277	839,277	839,277

Notes: Dependent variable: student math–reading performance difference (pooled z-score). Each column reports coefficients from a separate regression. Column (5) adds average teacher skill ($\bar{T}$) to test the stability of the teacher skill-difference coefficient. Column (6) adds average teacher skill to the specification with all controls. The instructional-time difference enters unstandardized, in hours per week. PISA-wave fixed effects absorbed. Weights rescaled so each country×wave cell receives equal weight. Standard errors clustered by country ($G = 25$) in parentheses. Stars based on analytical p-values. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. *Data sources:* PIAAC Cycle 1 (2012, 2015), PIAAC Cycle 2 (2022, 2023) and PISA 2009, 2012, 2018 and 2022.

Table 4 turns to the standard within-student, across-subject design. The outcome is each student's standardized mathematics score minus standardized reading score; the key regressor is the standardized teacher numeracy score minus standardized teacher literacy score. Thus, the model asks whether countries and periods with a more numeracy-oriented teacher workforce also have students whose achievement is more mathematics-oriented.

Column (1) yields a teacher skill-difference coefficient of 0.187. Adding the parent skill difference and the instructional-time difference reduces the coefficient to 0.166, which remains significant at the 1 percent level. Because the four underlying domains are standardized separately before differencing, while the differences themselves are not re-standardized, the coefficient is a relative-profile coefficient rather than a conventional standard-deviation effect: a one-unit shift in the teacher numeracy-minus-literacy profile is associated with a 0.166-unit shift in the student mathematics-minus-reading profile.

Columns (5) and (6) relax the symmetry restriction by adding average teacher skill. The coefficient on average teacher skill, ρ, is about 0.05 and marginally significant, while the

teacher skill-difference coefficient is virtually unchanged. The cross-sectional relationship therefore contains some asymmetry between numeracy and literacy, but this does not account for the estimated relationship between relative teacher and student skill profiles.

C Within Countries: Controlling for Persistent Subject-Specific Effects

Table 5: Within-Student Skill-Difference Model with Country Fixed Effects

	(1) Teacher skill diff.	(2) Parent skill diff.	(3) Instr.-time diff.	(4) All diff.	(5) Teacher skill diff. + $\bar{T}$	(6) All diff. + $\bar{T}$
Teacher skill difference	0.070*			0.075*	0.070*	0.077**
	(0.037)			(0.037)	(0.036)	(0.035)
Parent skill difference		0.017*		0.015*		0.015*
		(0.008)		(0.009)		(0.008)
Instructional-time difference			0.071	0.126		0.131
			(0.070)	(0.075)		(0.081)
$\bar{T}$					−0.004	−0.008
					(0.025)	(0.026)
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Overall R^2	0.07	0.06	0.06	0.07	0.07	0.07
Observations	839,277	839,277	839,277	839,277	839,277	839,277

Notes: Dependent variable: student math–reading performance difference (pooled z-score). Each column reports coefficients from a separate regression. Column (5) adds average teacher skill ($\bar{T}$) to test the stability of the teacher skill-difference coefficient. Column (6) adds average teacher skill to the specification with all controls. The instructional-time difference enters unstandardized, in hours per week. PISA-wave and country fixed effects absorbed. Weights rescaled so each country×wave cell receives equal weight. Standard errors clustered by country ($G = 25$) in parentheses. Stars based on analytical p-values. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. *Data sources:* PIAAC Cycle 1 (2012, 2015), PIAAC Cycle 2 (2022, 2023) and PISA 2009, 2012, 2018 and 2022.

Table 5 presents the paper's central test. Adding country fixed effects to the skill-difference model absorbs every time-invariant determinant of a country's mathematics–reading achievement gap. The coefficient is therefore identified by whether countries whose teacher workforces shift toward numeracy relative to literacy also experience a relative shift in student achievement toward mathematics.

In the current richest specification, the teacher skill-difference coefficient is 0.077. This is roughly half the pooled skill-difference estimate of 0.166. The decline is substantively

important: persistent country-by-subject factors account for a meaningful share of the conventional cross-subject association. But the coefficient remains positive and statistically significant under conventional country-clustered inference when these factors are removed. We interpret this as a reassuring result for the broader within-student, across-subject literature. The design's subject-specific omitted-variable concern matters for the magnitude of the coefficient, yet it does not explain the entire teacher-skill/student-achievement relationship.

The result is also notable given the measurement problem created by the stronger design. Teacher skills are estimated from finite PIAAC samples, and the identifying variation now comes from changes in a numeracy-minus-literacy difference. Classical sampling error is therefore amplified by differencing across domains and time and should attenuate the within-country coefficient. Our empirical-Bayes diagnostics indicate that the underlying country-level teacher means are measured with high reliability, but the long-difference estimate remains the specification in which attenuation is most likely to matter. For this reason, we regard 0.077 as a conservative estimate under classical measurement error and emphasize the persistence of the positive relationship rather than the precise magnitude.

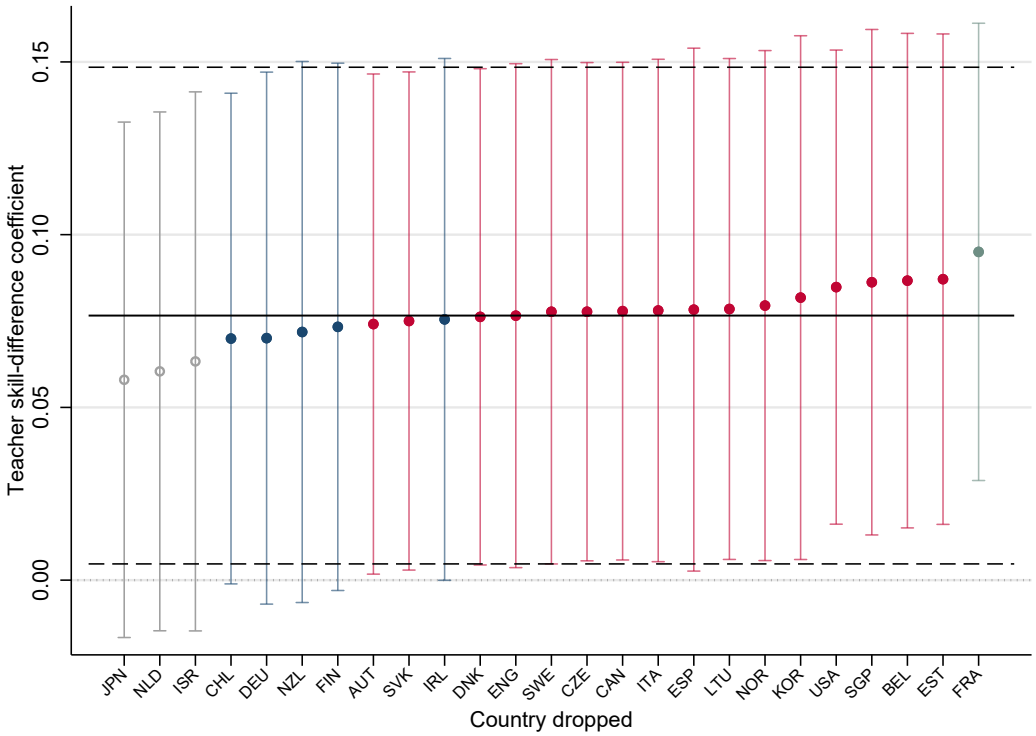
The symmetry diagnostic also improves once persistent country differences are absorbed. The coefficient on average teacher skill is -0.004 in column (5) and -0.008 in column (6), and statistically insignificant in both. Hence, the asymmetry observed in the pooled cross-country specification appears to arise primarily from between-country differences. Within countries over time, the parsimonious restriction that the relative numeracy and literacy effects are symmetric is not rejected.

The between–within decomposition sharpens this interpretation. The estimated between-country coefficient is $\beta_B = 0.218$, whereas the within-country coefficient is $\beta_W = 0.077$; both are statistically significant at the 5 percent level using conventional clustered inference. The pooled coefficient of 0.166 is therefore not a "between" estimate: it combines a strong between-country relationship with a smaller but still positive within-country relationship. The decomposition makes explicit what the repeated PIAAC data add to the earlier cross-sectional evidence.

D Robustness and Falsification

Influential countries. Figure 5 sequentially drops each of the 25 countries and re-estimates the country-fixed-effects specification. The teacher skill-difference coefficient remains positive in every leave-one-out sample. Its precision varies, as expected in a design identified by a limited number of country-level changes, but no single country accounts for the positive relationship.

Figure 5: Leave-One-Out Estimates of the Within-Country Teacher Skill Effect



Notes: The figure reports the teacher skill-difference coefficient from the country-fixed-effects specification, re-estimated after excluding each country in turn. The solid horizontal line is the full-sample coefficient; dashed lines are its 95% confidence interval. All regressions include country and PISA-wave fixed effects and control for average teacher skill, the instructional-time difference, and the parent skill difference. Standard errors are clustered by country. *Data sources:* PIAAC Cycles 1 and 2 and PISA 2009, 2012, 2018, and 2022.

Sign-based specification. Table 6 replaces the continuous skill differences with indicators for their sign. The dependent variable indicates whether a student is mathematics-dominant, and the key regressor indicates whether the teacher workforce is numeracy-dominant. The estimates have the same positive pattern as the continuous specification. This exercise is useful because it does not rely on the cardinal magnitude of the underlying skill differences.

Table 6: Binary Skill-Difference Model: Median Teacher Skills (Linear Probability Model)

	(1) Teacher skill-diff.	(2) Parent skill-diff.	(3) Time diff.	(4) All diff.
Teacher numeracy-dominant	0.038*			0.039*
	(0.020)			(0.021)
Parent numeracy-dominant		0.009		0.009
		(0.007)		(0.007)
Math instructional-time-dominant			−0.014	0.006
			(0.034)	(0.031)
Wave FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Overall R^2	0.04	0.04	0.04	0.04
Observations	839,277	839,277	839,277	839,277

Notes: This table reports linear probability model estimates of a non-parametric robustness check for the within-student skill-difference model with country fixed effects. The dependent variable equals one if a student's standardized math score exceeds the standardized reading score. Teacher numeracy-dominant and parent numeracy-dominant equal one if the corresponding standardized numeracy score exceeds the standardized literacy score. Math instructional-time-dominant equals one if weekly math instructional time exceeds weekly reading instructional time. All specifications include PISA-wave and country fixed effects. Weights are rescaled so that each country×wave cell receives equal weight. Standard errors clustered by country ($G = 25$) are reported in parentheses. Stars are based on analytical p-values: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. *Data sources:* PIAAC Cycles 1 and 2 and PISA 2009, 2012, 2018, and 2022.

National skill environment. A particularly important remaining concern is that the teacher skill profile may simply proxy for a broader within-country shift toward numeracy or literacy. We therefore construct the numeracy-minus-literacy skill profile of nonteacher adults and add it to the country-fixed-effects model. The most informative specification should include this adult skill difference jointly with average teacher skill, the predicted parent skill difference, and the instructional-time difference. Stability of the teacher coefficient in this horse race would distinguish teacher-specific changes from broader changes in national comparative cognitive advantage.

Occupation placebos. The most direct falsification test replaces the teacher numeracy–literacy difference with corresponding differences for other occupations. If the preferred

estimate reflects a general national shift in comparative skills, similar coefficients should appear for managers and other professional occupations that have no direct teaching role. We will complement these occupation-level placebos with matched nonteacher groups that reproduce teachers' age, gender, education, and sample-size composition. Matching is particularly useful because it makes the amount of sampling noise in the placebo skill measures comparable to that in the teacher measure.

Further specification and measurement checks. Several additional exercises belong in the appendix. First, estimate the unrestricted two-domain system allowing teacher numeracy and literacy to have separate aligned- and cross-domain effects. This shows whether the relative-profile coefficient is driven by the expected aligned-domain relationships rather than unusual cross-domain spillovers. Second, vary the PISA–PIAAC temporal alignment and report the already-implemented specification excluding PISA 2022, which is especially useful because language instructional time is not measured comparably in that wave. Third, summarize measurement robustness in one table: median versus mean teacher skills, empirical-Bayes-shrunk means (and, when available, the skill difference itself), composition-adjusted teacher skills, a restriction to country-cycles with larger teacher samples, and a broader teacher definition including kindergarten teachers. Finally, report the fixed-reference-rank specification as a scale-free robustness check.

VI Teacher Skills and Adult Skill Formation

PISA captures student achievement near the end of compulsory schooling. As complementary evidence, we ask whether changes in teacher skill profiles are also mirrored in the skills of younger adults observed in PIAAC. We focus on adults aged 25–34, whose schooling is more proximate to the teacher cohorts represented in the two PIAAC cycles than that of older adults. For country c and cycle s, let $S^N_{c,s}$ and $S^L_{c,s}$ denote median numeracy and literacy, and define $D^S_{c,s} = S^N_{c,s} - S^L_{c,s}$. We estimate

$$\Delta D^S_c = \phi \Delta D^T_c + \Delta X_c' \omega + \xi_c, \tag{23}$$

where $\Delta(\cdot)$ denotes the change between PIAAC Cycles 1 and 2. The vector ΔX_c captures compositional changes in the young-adult population, including migration and educational attainment.

The coefficient ϕ asks whether countries in which the teacher workforce shifts toward numeracy relative to literacy also exhibit a corresponding shift among younger adults. This exercise is necessarily more descriptive than the PISA analysis. We cannot link individual

adults to the teachers who taught them, and exposure timing is less precise. Its value is therefore complementary: it examines whether the same cross-domain, within-country pattern appears in an outcome measured independently of PISA and at a later age.

VII Determinants of Teacher Cognitive Skills over Time

The descriptive evidence shows substantial within-country movement in both teacher skills and teachers' relative position in the tertiary-educated workforce. We next ask which labor-market conditions are associated with these changes. The objective here is not to claim causal effects from a 25-country panel, but to distinguish mechanisms that are consistent with changing selection into teaching from purely mechanical population-skill trends.

We relate changes in teacher skills between PIAAC cycles to changes in national teacher labor-market conditions:

$$\Delta T_c^j = \Delta W_c' \pi + \Delta Z_c^{j\prime} \psi + \nu_c^j, \tag{24}$$

where T_c^j is a statistic of the teacher skill distribution in domain $j \in \{N, L\}$. Outcomes include changes in mean or median skills, teachers' percentile position among tertiary-educated nonteachers, and—where sampling precision permits—distributional measures such as selected percentiles.

The vector ΔW_c contains country-wide changes that may alter recruitment and retention in teaching: relative teacher pay, aggregate labor-market conditions, women's access to high-skill occupations outside teaching, and measures of teacher status. These mechanisms follow naturally from evidence that outside opportunities and compensation affect occupational selection into teaching (Corcoran, Evans and Schwab, 2004; Bacolod, 2007; Nagler, Piopiunik and West, 2020; Biasi, 2021; Estrada and Lombardi, 2023). The vector ΔZ_c^j captures domain-specific changes, such as differential labor-market returns to quantitative versus verbal skills or relative shortages of mathematics and language teachers.

Because equation (24) is estimated in first differences, all stable country characteristics are removed. The remaining variation is nevertheless observational and limited to a small number of countries. We therefore interpret the estimates as descriptive evidence on mechanisms and emphasize magnitudes, robustness to influential countries, and consistency across alternative measures of teacher selection rather than conventional significance alone.

A Alternative Professional Opportunities for Women

Analysis of changes in women's professional opportunities and their association with teacher selection will be added here.

B Teacher Pay, Returns to Skills, and Occupational Selection

This subsection will examine whether changes in teacher wage premiums and in the returns to cognitive skills inside and outside teaching are associated with changes in the skill composition of the profession. The repeated design extends cross-sectional evidence on teacher pay and skills by focusing on changes within countries.

VIII Conclusions

Teacher cognitive skills are not a fixed feature of national education systems. Across 25 countries observed in two PIAAC cycles, average teacher skills move little, but individual countries experience sizable gains and losses, and teachers' position within the tertiary-educated skill distribution shifts markedly. These facts open a new margin for studying teacher quality: not only why countries differ in the skills of their teachers, but why the skill composition of teaching changes within countries over time.

The repeated data also strengthen the evidence linking teacher skills to student achievement. The standard within-student, across-subject model yields a teacher skill-difference coefficient of 0.166. Once country fixed effects absorb persistent country-specific determinants of the mathematics–reading gap, the coefficient falls to 0.077 but remains positive and statistically significant. The between–within decomposition tells the same story: the cross-country association is substantially stronger than the within-country association, but the latter does not disappear. Persistent subject-specific heterogeneity therefore matters for the magnitude of conventional cross-subject estimates, yet it does not explain the entire relationship.

This result is useful beyond the present application. Within-student, across-subject designs are widely used precisely because they eliminate student heterogeneity and inputs common across subjects. Our repeated setting permits a direct assessment of an important residual concern—persistent subject-specific factors—that is usually untestable in a single cross section. Finding attenuation but not elimination is consequently a qualified validation of that design logic. Moreover, because the within-country estimate is based on changes in country-level teacher skill differences, classical sampling error should work toward attenuation. We therefore view the within-country coefficient as deliberately conservative and avoid interpreting its precise magnitude as a structural causal effect.

The paper contributes along three dimensions: it documents the international evolution of directly measured teacher cognitive skills and occupational selection; it adds a time dimension to cross-subject teacher-effect designs and shows that the teacher-skill relationship survives a substantially stronger comparison; and it uses repeated international data to study how teacher labor-market conditions are associated with changes in the skill composition of the profession. Completing the occupation-placebo, national-skill-environment, and teacher-labor-market analyses will further clarify how much of the remaining within-country relationship is specific to teachers and which institutional forces are associated with changes in teacher quality.

References

Araujo, M. Caridad, Pedro Carneiro, Yyannú Cruz-Aguayo, and Norbert Schady. 2016. "Teacher Quality and Learning Outcomes in Kindergarten." *Quarterly Journal of Economics*, 131(3): 1415–1453.

Bacher-Hicks, Andrew, and Cory Koedel. 2023. "Estimation and Interpretation of Teacher Value Added in Research Applications." In *Handbook of the Economics of Education*. Vol. 6, 93–134. Elsevier.

Bacolod, Marigee P. 2007. "Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality." *The Review of Economics and Statistics*, 89(4): 737–751.

Bardach, Lisa, and Robert M. Klassen. 2020. "Smart Teachers, Successful Students? A Systematic Review of the Literature on Teachers' Cognitive Abilities and Teacher Effectiveness." *Educational Research Review*, 30: 100312.

Barrios-Fernández, Andrés, and Marc Riudavets-Barcons. 2024. "Teacher Value-Added and Gender Gaps in Educational Outcomes." *Economics of Education Review*, 100: 102541.

Bau, Natalie, and Jishnu Das. 2020. "Teacher Value Added in a Low-Income Country." *American Economic Journal: Economic Policy*, 12(1): 62–96.

Bertoni, Eleonora, Gregory Elacqua, Carolina Mendez, and Humberto Santos. 2024. "Teacher Selection Instruments and Teacher Value-Added: Evidence From Peru." *Educational Evaluation and Policy Analysis*, 46(1): 53–81.

Biasi, Barbara. 2021. "The Labor Market for Teachers under Different Pay Schemes." *American Economic Journal: Economic Policy*, 13(3): 63–102.

Bietenbeck, Jan, Marc Piopiunik, and Simon Wiederhold. 2018. "Africa's skill tragedy: Does teachers' lack of knowledge lead to low student performance?" *Journal of Human Resources*, 53(3): 553–578.

Bietenbeck, Jan, Natalie Irmert, and Mohammad H Sepahvand. 2023. "Teacher quality and cross-country differences in learning in francophone Sub-Saharan Africa." *Economics of Education Review*, 96: 102437.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review*, 104(9): 2633–2679.

Clotfelter, Charles T, Helen F Ladd, and Jacob L Vigdor. 2010. "Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects." *Journal of Human Resources*, 45(3): 655–681.

Corcoran, Sean P, William N Evans, and Robert M Schwab. 2004. "Women, the labor market, and the declining relative quality of teachers." *Journal of Policy Analysis and Management*, 23(3): 449–470.

Dee, Thomas S. 2005. "A teacher like me: Does race, ethnicity, or gender matter?" *American Economic Review*, 95(2): 158–165.

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

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(65): 5–55.

Dolton, Peter, Oscar Marcenaro, Robert de Vries, and Po-Wen She. 2018. "Global teacher status index 2018."

Estrada, Ricardo, and María Lombardi. 2023. "Skills and Selection into Teaching: Evidence from Latin America." *Economía LACEA Journal*, 22(1): 218–239.

Golsteyn, Bart H. H., Stan Vermeulen, and Inge de Wolf. 2016. "Teacher Literacy and Numeracy Skills: International Evidence from PIAAC and ALL." *De Economist*, 164(4): 365–389.

Grönqvist, Erik, and Jonas Vlachos. 2016. "One size fits all? The effects of teachers' cognitive and social abilities on student achievement." *Labour Economics*, 42: 138–150.

Hanushek, Eric A, and Steven G Rivkin. 2006. "Teacher quality." *Handbook of the Economics of Education*, 2: 1051–1078.

Hanushek, Eric A, and Steven G Rivkin. 2012. "The distribution of teacher quality and implications for policy." *Annual Review of Economics*, 4(1): 131–157.

Hanushek, Eric A, Babs Jacobs, Guido Schwerdt, Rolf Van der Velden, Stan Vermeulen, and Simon Wiederhold. 2023. "Where do STEM graduates stem from? The intergenerational transmission of comparative skill advantages." National Bureau of Economic Research.

Hanushek, Eric A, Marc Piopiunik, and Simon Wiederhold. 2019. "The value of smarter teachers: International evidence on teacher cognitive skills and student performance." *Journal of Human Resources*, 54(4): 857–899.

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(2): 236–240.

Jackson, C. Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes." *Journal of Political Economy*, 126(5): 2072–2107.

Jackson, C Kirabo, Jonah E Rockoff, and Douglas O Staiger. 2014. "Teacher effects and teacher-related policies." *Annual Review of Economics*, 6(1): 801–825.

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

Leigh, Andrew, and Chris Ryan. 2008. "How and Why Has Teacher Quality Changed in Australia?" *Australian Economic Review*, 41(2): 141–159.

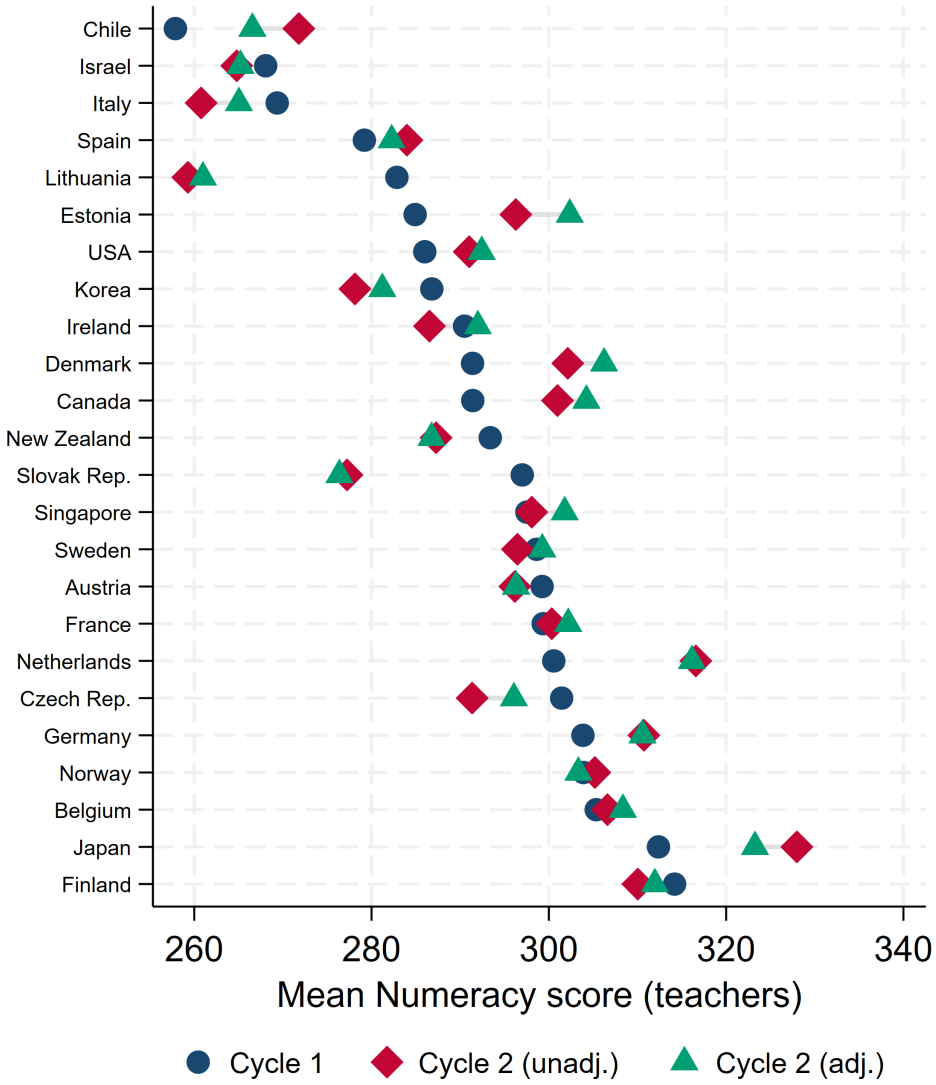
Metzler, Johannes, and Ludger Woessmann. 2012. "The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation." *Journal of Development Economics*, 99(2): 486–496.

Nagler, Markus, Marc Piopiunik, and Martin R West. 2020. "Weak markets, strong teachers: Recession at career start and teacher effectiveness." *Journal of Labor Economics*, 38(2): 453–500.

Walters, Christopher. 2024. "Empirical Bayes Methods in Labor Economics." *Handbook of Labor Economics*, 5: 183–260.

A Composition Adjustment of Teacher Skills

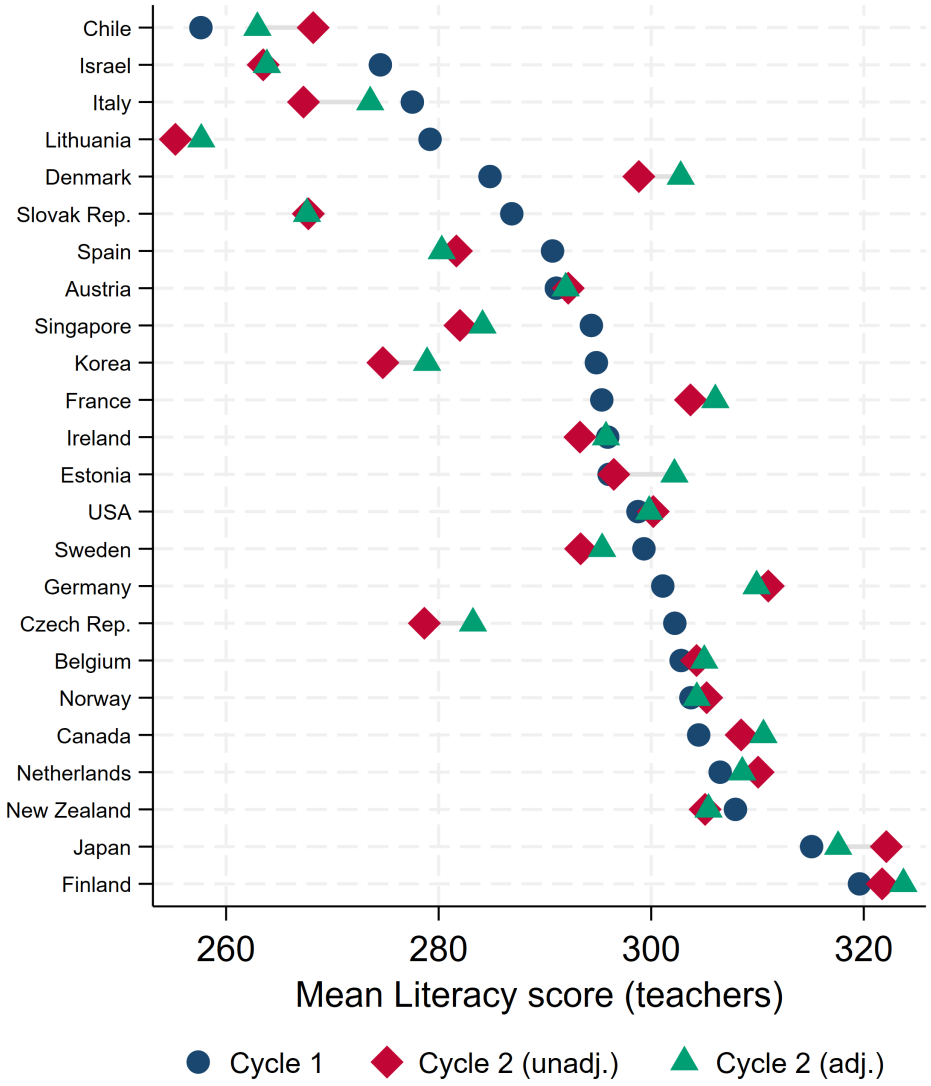
Figure A1: Composition Adjustment of Mean Teacher Skills: Numeracy



Data sources: PIAAC Cycle 1 (2012, 2015) and PIAAC Cycle 2 (2022, 2023).

Notes: The figure shows mean teacher numeracy by country for Cycle 1 (blue), for Cycle 2 without adjustment (red), and for Cycle 2 after reweighting the Cycle 2 teachers to the Cycle 1 composition by inverse probability weighting on age, gender and teacher type (green). The grey horizontal line indicates the size of the adjustment for each country. Countries are ordered by their Cycle 1 mean.

Figure A2: Composition Adjustment of Mean Teacher Skills: Literacy

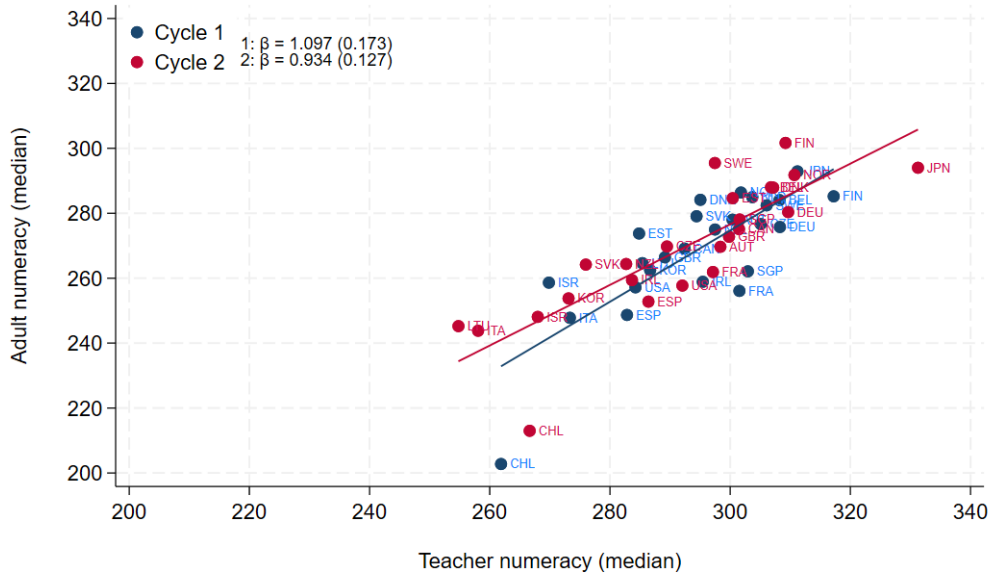


Data sources: PIAAC Cycle 1 (2012, 2015) and PIAAC Cycle 2 (2022, 2023).

Notes: The figure shows mean teacher literacy by country for Cycle 1 (blue), for Cycle 2 without adjustment (red), and for Cycle 2 after reweighting the Cycle 2 teachers to the Cycle 1 composition by inverse probability weighting on age, gender and teacher type (green). The grey horizontal line indicates the size of the adjustment for each country. Countries are ordered by their Cycle 1 mean.

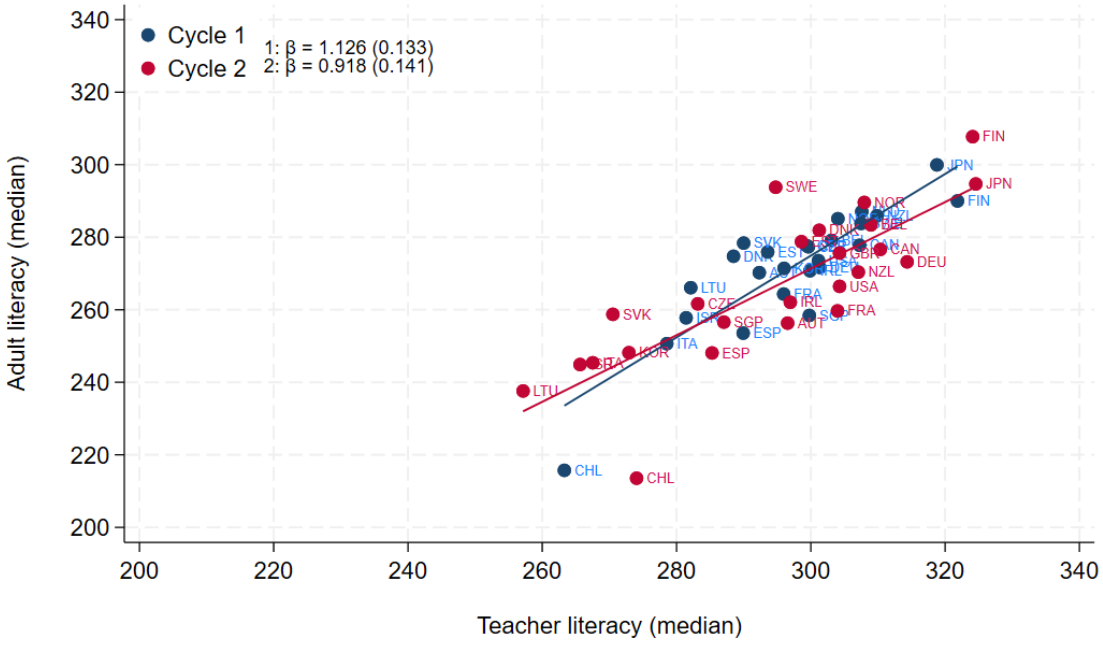
B Supplementary Figures

Figure B1: Teacher Cognitive Skills vs. Adult Cognitive Skills: Numeracy



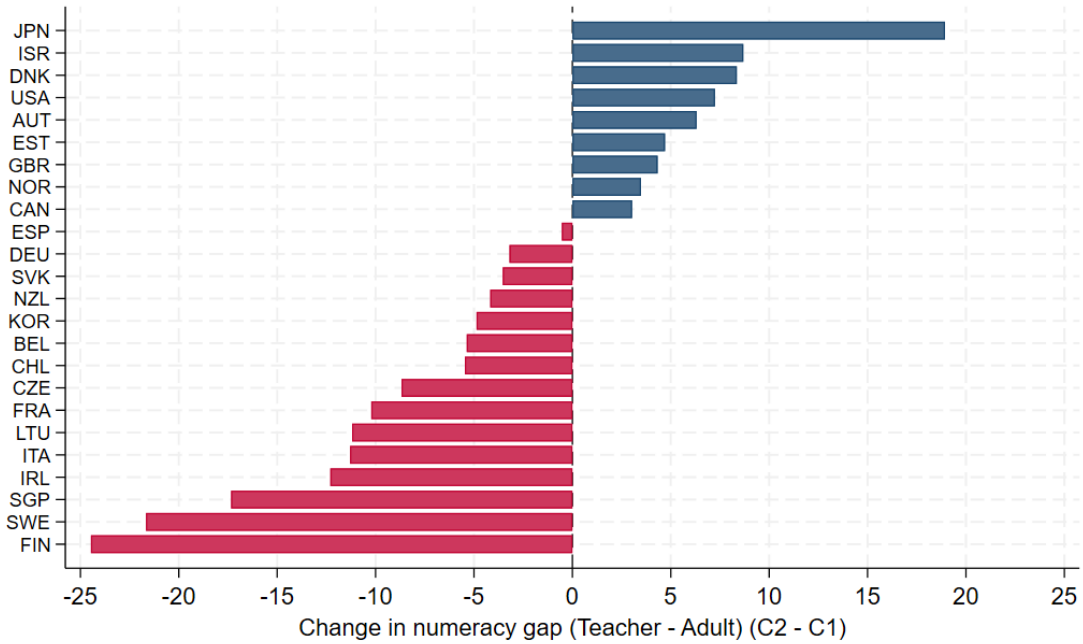
Data sources: PIAAC Cycle 1 (2012, 2015) and PIAAC Cycle 2 (2022, 2023).
Notes: The figure displays the relationship between median teacher numeracy skills and median adult numeracy skills for 25 countries across two PIAAC cycles. Adult skills in Cycle 2 are computed from Public Use Files (Netherlands excluded due to data availability). The fitted lines show the bivariate relationship between teacher and adult skills for each cycle separately.

Figure B2: Teacher Cognitive Skills vs. Adult Cognitive Skills: Literacy



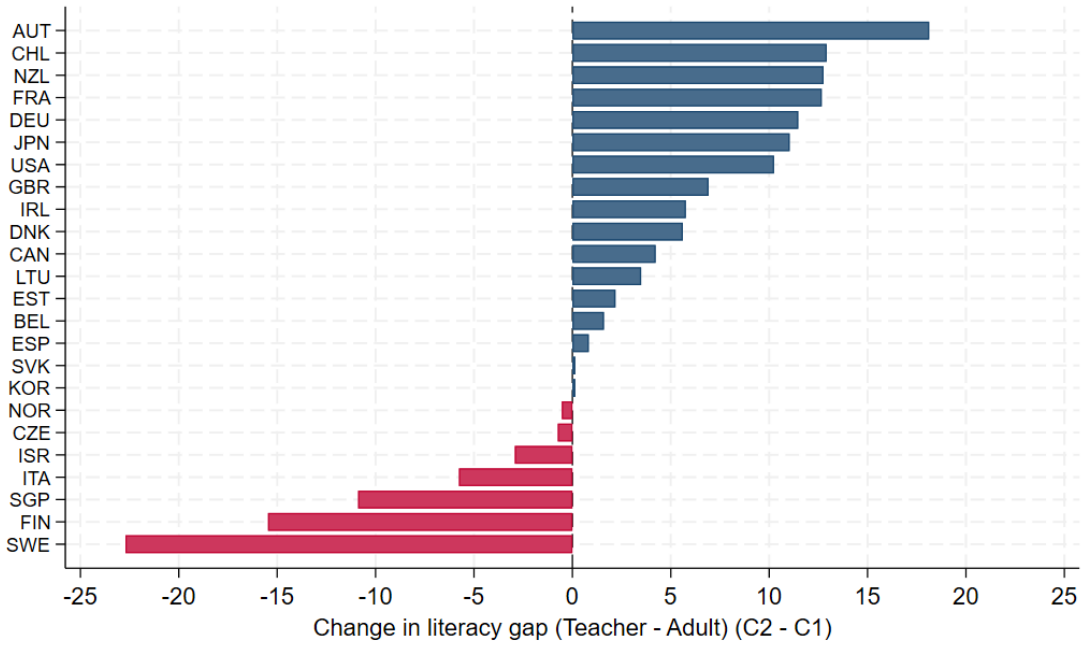
Data sources: PIAAC Cycle 1 (2012, 2015) and PIAAC Cycle 2 (2022, 2023).
Notes: The figure displays the relationship between median teacher literacy skills and median adult literacy skills for 25 countries across two PIAAC cycles. Adult skills in Cycle 2 are computed from Public Use Files (Netherlands excluded due to data availability). The fitted lines show the bivariate relationship between teacher and adult skills for each cycle separately.

Figure B3: Changes in Teacher-Adult Skill Difference: Numeracy



Data sources: PIAAC Cycle 1 (2012, 2015) and PIAAC Cycle 2 (2022, 2023).
Notes: The figure displays changes in the difference between median teacher numeracy skills and median adult numeracy skills between Cycle 1 and Cycle 2. Positive values indicate that teacher skills increased more (or decreased less) than adult skills. Negative values indicate that teacher skills increased less (or decreased more) than adult skills. Adult skills in Cycle 2 are computed from Public Use Files (Netherlands excluded due to data availability).

Figure B4: Changes in Teacher-Adult Skill Difference: Literacy



Data sources: PIAAC Cycle 1 (2012, 2015) and PIAAC Cycle 2 (2022, 2023).
Notes: The figure displays changes in the difference between median teacher literacy skills and median adult literacy skills between Cycle 1 and Cycle 2. Positive values indicate that teacher skills increased more (or decreased less) than adult skills. Negative values indicate that teacher skills increased less (or decreased more) than adult skills. Adult skills in Cycle 2 are computed from Public Use Files (Netherlands excluded due to data availability).