The Effect of Teachers' On-the-job Training on Student Achievement

Lund University

Tilman Bretschneider, Ida Haggren, Benjamin Maday & Christina Maschmann

Abstract

We study the effect of teachers' on-the-job training on educational achievement. Using cross-sectional data from the Programme for International Student Assessment (PISA), we exploit a within-student fixed effects approach and show that professional development for teachers does significantly improve students' standardised test scores. More specifically, teachers' participation in on-the-job training increases the test scores by roughly 0.013 of a standard deviation. Our analysis further shows that especially students with migration background seem to benefit from such career interventions.

Key Words: PISA, On-the-job Teacher Training, Education.

Econometric Game 2024
April 19, 2024

1 Introduction

How can teachers become more effective in a way which would benefit students? This question is widely discussed in the economic literature, especially in the aftermath of the COVID-19 pandemic. In the past few years, policymakers in different countries have introduced several policies to improve teaching quality and consequently, student performance. As teachers usually do not switch professions during their career, on-the-job training is particularly important to provide them with new insights on pedagogical methods or new developments in their field. Therefore, a large number of countries in the OECD increasingly support teachers by introducing new opportunities and extensive measures for formal professional development (OECD, 2020). For example, in 2014 the Ministry of Education and Culture in Finland funded a new on-the-job training program for teachers, aiming to further educate 50,000 trained teachers over two years (Mullis, 2020). Despite the societal acknowledgement that participation in on-the-job training will tend to increase teachers' skills and their confidence in their knowledge (OECD, 2020), the economic literature on education provides only little consensus on whether this professional development may benefit student performance.

We contribute to this rather scarce literature and test whether teachers' participation in on-the-job training increases students' standardised test scores. For our empirical analysis we draw data from PISA (Programme for International Student Assessment). This database contains mathematics, science and reading test scores for 15-year-old students in OECD countries. Applying a within-student between-subject fixed effects approach, we focus on the 2022 test scores and the following six OECD countries: Australia, Columnbia, Costa Rica, Germany, South Korea and Portugal. Our main results suggest that on-the-job training for teachers significantly increases students' achievement by roughly 0.013 of a standard deviation. Furthermore, our heterogeneity analysis shows that students with migration background benefit most from on-the-job training for teachers as perhaps teachers learn how to address those children's needs and can unlock their unrealised potential. Our analysis suggest that the effect of on-the-job training works through better communication with parents and more tailoring to the specific students, although these results cannot be interpreted casually.

This paper is organized as follows. Section 2 reviews the already available literature of effects of teachers' on-the-job training on student performance. Section 3 describes the data used in our empirical analysis as well as details on our sample selection and existing trends within and across the OECD countries. Section 4 presents our empirical strategy and Section 5 outlines our main results as well as heterogeneity analyses. Section 6 discusses possible mechanisms and several robustness checks. Finally, Section 7 contains the conclusions.

2 Literature Review

There exists two strands of literature that investigate teachers' contribution to students' performance. Hereby, the evidence emphasizes that teachers' effectiveness is an important driver of students' outcomes.

The first strand focuses on a value-added approach and estimates the overall contribution of teachers to students' achievements using longitudinal data. This approach suggests that a teacher's added value can be viewed as a proxy for a teacher's motivation to increase students' human capital (Chetty et al., 2014; Jackson et al., 2014). Notably, this effect is rather unobserved and is not depicted by observable inputs into the educational production function (Jackson et al., 2014). The value-added approach uses least square regressions and estimates a teacher's relative productivity relying on test score variations across students linked to an identical teacher. Hence, it does not identify which specific teacher characteristics are important for teaching quality. However, the literature in this strand does agree that there is an important variation in teacher effectiveness. Chetty et al. (2014) find that students being matched with teachers who indicate a high added value are more likely to attend college, benefit more on the labor market and tend to live in a higher quality neighborhood as adults.

Despite providing nearly unbiased estimates of teachers' effect on student performance, the value-added approach does not investigate which teachers' characteristics are specifically important to increase teachers' effectiveness. This information is particularly important for policymakers and educational institutions. In the process of teacher recruitment, schools usually only observe teachers' characteristics, such as teaching experience and educational background rather than their added value. Those observable characteristics can help to establish a minimum level of teaching quality. Furthermore, a sophisticated understanding of how certain teacher-specific characteristics affect students' performance can assist policymakers to re-assess prevailing policies on teacher recruitment or career incentives (Clotfelter et al., 2010).

Hence, the second strand of literature pays attention to the effects of teachers' observable characteristics on students' achievements. Most of the studies focus on factors determining teachers' wages, such as level of education or teaching experience (Clotfelter et al., 2010; Hanushek et al., 2013). To the best of our knowledge, there is little consensus and only limited evidence whether professional development may benefit student performance. The existing studies find either positive effects (see Angrist and Lavy (2001) and Bressoux et al. (2009)) or no effects at all (Harris & Sass, 2011; Jacob & Lefgren, 2004).

3 Data and Descriptive Evidence

3.1 PISA Database

We obtain our data on students' performance in mathematics and science from the Programme for International Student Assessment (PISA). PISA was first conducted in 2000 by the OECD and the data collection has since been repeated every three years, with the exception of 2021, due to the COVID-19 pandemic. The number of participating countries in every sampling period is more than 50 developing and developed countries. In each wave, PISA randomly selects nationally representative samples of 15-year-old students and assesses them using standardised multiple-choice tests in science, mathematics and reading. The main goal of these tests is to measure students' knowledge of the subject to obtain internationally comparable data on their educational achievement. This can be leveraged to improve educational policies, human capital accumulation and to resolve inequalities (Bietenbeck & Collins, 2023; Hanushek et al., 2013).

PISA employs a two-stage sampling design in most of the countries. First, a random sample of schools in which 15-year-old students are enrolled, is drawn. Hereby, the probability for a school being selected is proportional to its size i.e., the estimated number of 15-year-old students attending. In the second stage, PISA samples 35 students of the 15-year-students from the eligible schools where each student has the same sampling probability (Hanushek et al., 2013).

The test scores are standardised to have a mean of 500 and a standard deviation of 100 across all OECD countries participating in PISA 2000. PISA uses a method of plausible values to account for the uncertainty of reported statistics resulting from differences in the distributed tasks. Hereby, ten plausible values are provided for each subject. For the sake of simplicity, we use the average test score per student for each subject as our main outcome variable[1].

The test score results are accompanied by a rich set of background information on each student and school. Students are asked to fill out a questionnaire on their demographic characteristics but also their family background. Also teachers answer a questionnaire containing information on their educational background, experience and work environment. Furthermore, school principals are asked to provide information on a school's institutional settings and its endowments, such as admittance policies, school's location or its digital resources.

3.2 Sample Selection and Descriptive Statistics

In our analysis, we focus on the 2022 test results of the following six OECD countries: Australia, Columbia, Costa Rica, Germany, South Korea, and Portugal[2]. We consider a pooled sample containing 31,354 observations of 15-year-old students taught by 3892 teachers in 1512 schools.

To draw our final sample, we omit all observations where we could not successfully link students to their math or science teachers. Furthermore, it is important for our identification strategy to relate students to exactly one teacher per subject as violating this condition erases the unique assignment of teacher as well as class characteristics. Moreover, we drop all student observations which have the same teacher in both subjects as those would introduce a lack of within-student between-subject variation.

We define our treatment variable for on-the-job training as a binary indicator that takes the value of 1 if a teacher answers "yes" to the question "*Are you required to take part in professional development activities?*", and 0 otherwise. We argue that this is a suitable measure for teachers' on-the-job training as it is clearly distinguishable from formal teacher education but also informal on-the-job training by working experience. We present descriptive statistics of the treatment variable for each country in our sample in.

A part of our identification strategy is relying on a rich set of observed teacher characteristics. Specifically, we use teacher's gender, age, formal education, and work experience as a teacher. We categorise most control variables into groups to allow for nonlinear effects. For age, the groups are below 30, 30–39, 40–49, and above 50. Experience is grouped into 0–2, 3–5, and 6 or more years of experience.

Table 5 presents the mean, median, and standard deviation for the age and professional experience of teachers and for the mathematics and science test scores of their pupils in the selected countries. It also includes the ratio of teachers who have acquired at least a master's degree, and the ratio of those who have received some kind of on-the-job training according to

[1] An alternative approach would be to use jackknife repeated replication technique as a resampling method.

[2] We limit our analysis to these specific countries as only those allow us to conduct a clear matching procedure of students and teachers.

the survey question discussed above. The selected nations show big differences in the number of both on-the-job and university-educated instructors, presumably due to the different legal requirements toward teachers, and it is not clear from the metrics which approach produces the best results.

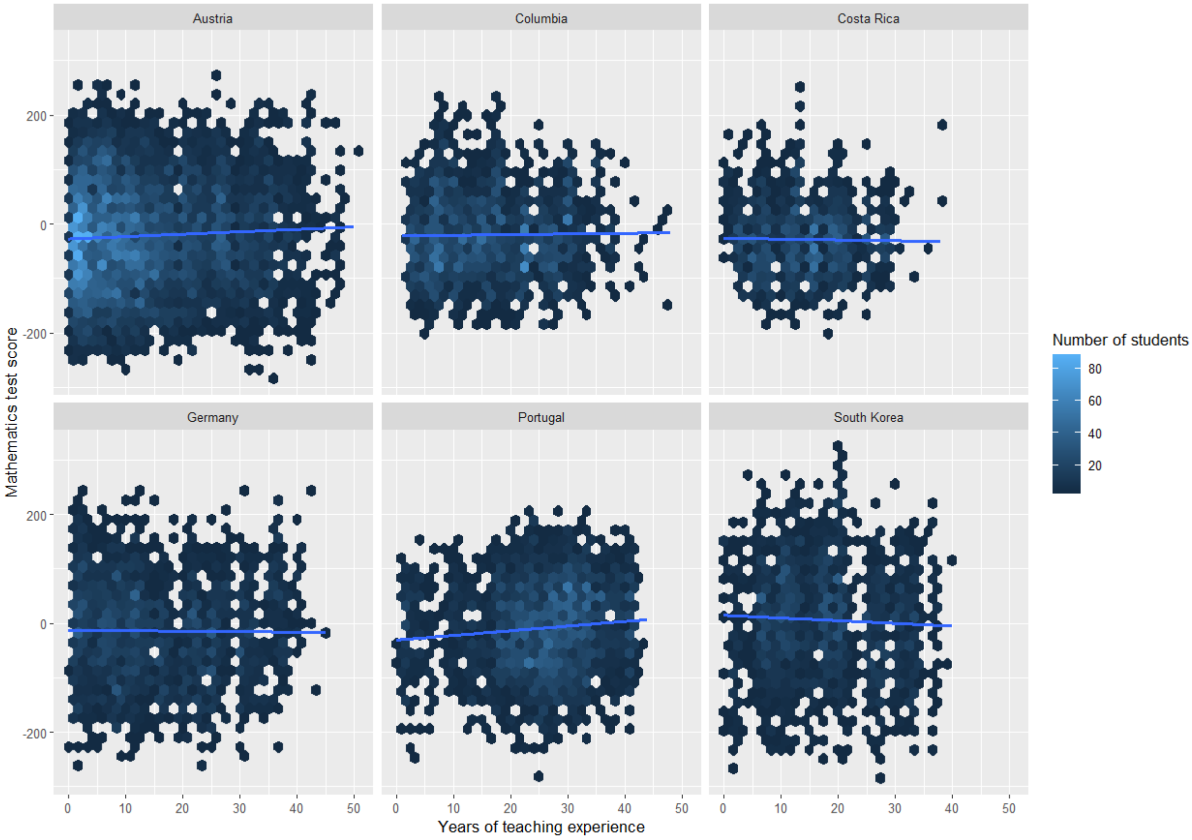

Figure 1: Demeaned mathematics test scores of students grouped by country in respect to years of teaching experience of teachers for the PISA wave 2022.

Figure 1 and Figure 2 show the relationship between students' performance on the mathematics and science tests respectively in relation to their teachers years of professional experience by country. The lighter a hex is, the more students are present in that subsection of the graph. We can observe that there is only a small correlation between the overall experience of tutors and the results of their pupils, and the effect is not consistent between countries. In Austria and Portugal, the teacher's experience increased a given student's results while in South Korea and Costa Rica, it decreased. These country-specific patterns however are consistent between subjects, suggesting an underlying structural reason due to perhaps the teacher allocation system in a given country.

4 Empirical Strategy

Analysing the effect of on-the-job training of teachers on student achievement using a non-experimental approach will potentially result in biased estimates due to endogeneity concerns. The main reason is that the assignment of teachers and students to classrooms is usually driven by self-selection rather than random assignment mechanisms (see e.g. Bietenbeck, 2014; Lavy,

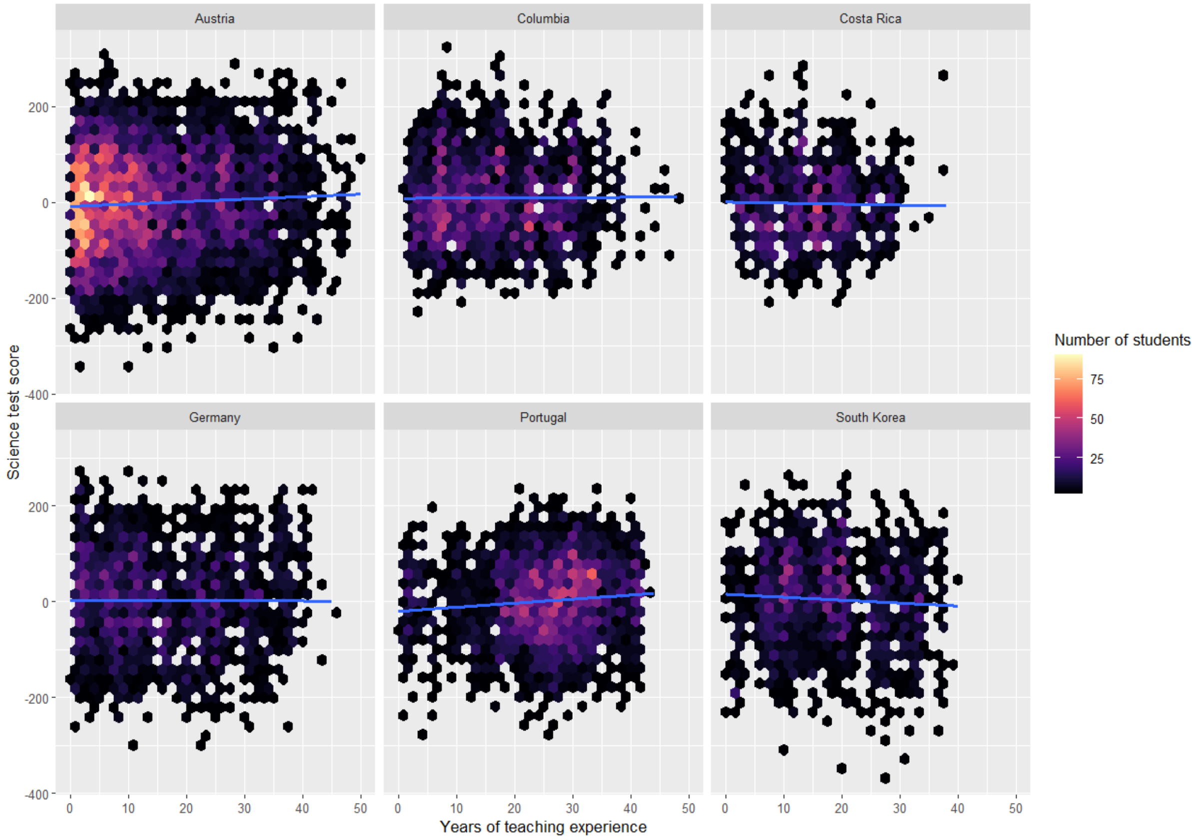

Figure 2: Demeaned science test scores of students grouped by country in respect to years of teaching experience of teachers for the PISA wave 2022.

2015). Hereby, the following confounders are raising our concerns. First, students and their parents might have particular preferences for specific teacher characteristics and hence, sort themselves into schools or classrooms with a focus on formal on-the-job teacher training. This selection would be problematic if, for instance, students with low unobserved ability might prefer to attend schools in which teachers tend to invest in their professional development more frequently. This would result in a downward bias of the true estimated effect of professional development of teachers on student achievements. Second, we claim that teachers' participation in on-the-job training is potentially correlated with unobserved teacher characteristics, such as their motivation or their ability. For example, teachers might adjust their investment of their professional development to the students in their classroom. Furthermore, motivated teachers who are passionate about their profession might have strong preferences for an ongoing professional development in order to increase their network or to learn new skills. Third, more resourceful schools might attract students from higher socioeconomic backgrounds while investing more into on-the-job training of their teachers which would lead to an upward bias of estimates omitting these factors. We actually observe that students from higher socioeconomic status parents have on average a higher propensity to be taught by a teacher that receives on-the-job training, see Figure 4.

To tackle those concerns, we follow the recent literature in economics of education and exploit an empirical strategy based on within-student between-subject variation. We introduce student fixed effects which helps to hold observed and unobserved students' characteristics, such as race, motivation, ability, family background and country of residence constant. This identification strategy follows essentially a panel data approach, requiring each student to be observed at least twice. Instead of observing a student in two different points in time, the same student should be observed in two different subjects but at the same point in time (Bietenbeck, 2014; Clotfelter et al., 2010; Dee, 2007; Lavy, 2015). The PISA data allows us to exploit this feature for the subjects mathematics and science. Therefore, we follow this approach and formulate a within-student fixed effect model which additionally allows the outcome variable, i.e. overall students' achievement, to be a function of observable teacher characteristics. Hence, in our main empirical approach, we focus on the following educational production function:

$$S_{ijk} = \alpha + \beta Training_{ijk} + \mathbf{T}_{jk}\gamma + \delta_i + \epsilon_{ijk} \quad (1)$$

with S_{ijk} being student i's standardised test score in subject k taught by teacher j. $Training_{ijk}$ is our treatment variable for on-the-job training of a teacher. Consequently, β is our parameter of interest. $\mathbf{T}_{jk}$ is a vector of observable teacher characteristics and δ_i represents the student fixed effects, i.e. observable and unobservable subject-invariant student characteristics. Finally, ϵ_{ijk} is the unobserved error term. One advantage of this approach is that the student fixed-effects δ_i also absorb subject-invariant school fixed effects, as we are able to observe the same student in the same school which addresses the third potential confounding factor. Therefore, it also controls for any school-specific factors which might influence the standardised test scores (Bietenbeck, 2014). This allows us to take into account all observable and unobservable subject-invariant variations within students and schools. Introducing controls for teacher characteristics further captures differences between teacher attributes for the same student in both subjects.

Our key identifying assumption is that the error term is uncorrelated with the treatment variable, i.e. on-the-job training of teachers. To ensure that this assumption holds, we address the following concerns. First, there might exist subject-specific unobserved characteristics which might influence the test scores but are also correlated with our treatment variable. For example, teachers might sort themselves in one of the subjects or adjust the amount of on-the-job training according to the subject they are teaching. This might lead to a selection bias and hence, jeopardize our results. However, even though there might be a certain variation between mathematics and science, we argue that those subjects demand very similar cognitive skills and student motivation. Furthermore, both subjects probably require a similar level of on-the-job training for teachers. This mitigates our concerns for future analysis.

Second, unobserved teachers' motivation might lead to a higher take-up rate of professional development courses. This leads to a potential correlation between unobserved teacher characteristics and our treatment variable, inducing additional channels on students' performance beside the actual treatment. This will lead to an omitted variable problem and an overestimation of the actual effect. We follow the current literature and – besides controlling for a rich set of observed teacher characteristics – check the coefficient stability by investigating the limit of this omitted variable bias. For that, we apply Oster (2019)'s approach on unobservable selection as a robustness check to our analysis.

Table 1: Estimated Effects of On-the-job Teacher Training on Standardized Student Test Scores

	(1) OLS	(2) FE	(3) FE
On-the-job Training (1=yes)	0.182***	0.0131**	0.0128**
	(0.011)	(0.0059)	(0.0060)
Teacher Controls	YES	NO	YES
Student Fixed Effects	NO	YES	YES
N	85,059	80,648	80,648

Notes: Results from weighted student fixed effects regressions with student standardized test scores as dependent variable. Test scores are standardized with mean zero and standard deviation of one. All regressions include subject fixed effects and a control for teachers' subject-specific specialization degree in either math or science. PISA sampling weights are utilized. Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

5 Results

5.1 Effects of Teachers' On-the-Job Training on Overall Student Achievement

As a preliminary analysis, we estimate the effects of on-the-job training of teachers on students' performance by running a naive OLS regression, including controls for teacher characteristics. We find a positive association between a teacher's professional development and students' test scores (column (1) of Table 1). The coefficient is highly statistically significant (1%-level) and fairly large, indicating on average approximately 0.18 standard deviations higher test scores for students whose teachers are required to do on-the-job training. As mentioned in Section 4, using such a non-experimental approach will result in biased estimates due to selection of students and teachers and hence should be considered with caution.

Columns (2) and (3) of Table 1 present the estimates using our main identification strategy that uses student fixed effects as we are able to exploit the fact that we observe students test scores in mathematics and science. The magnitude of these estimates is much smaller compared to the OLS estimates, indicating on average a 0.013 standard deviations increase in test scores when teachers receive on-the-job education. They are furthermore robust to the inclusion of teacher characteristics as a control variables and both statistically significant at the 5%-level. The fixed effects estimate provide strong evidence that OLS estimates suffer from endogeneity issues. As estimates are more than ten times larger, it seems plausible that the main driver behind them is selection of better performing students into better schools that provide better training for their teachers.

5.2 Heterogeneity Analysis

We evaluate the heterogeneity of our treatment effect along two dimensions: migration status and parents' education level. These results are found in Table 2.

Between the students who are first- and second-generation migrants (column (3)), and those

Table 2: Heterogeneity analysis of results by migration status and socioeconomic status

	(1) Main	(2) Native	(3) Migrant	(4) High Ed.	(5) Low Ed.
On-the-job Training (1=yes)	0.0128**	0.00700	0.0328**	0.0174**	0.00777
	(0.00598)	(0.00627)	(0.0166)	(0.00778)	(0.00933)
Teacher Controls	YES	YES	YES	YES	YES
Student Fixed Effects	YES	YES	YES	YES	YES
N	80,648	64,738	15,910	54,826	25,822

Notes: Results from weighted student fixed effects regressions with student standardized test scores as dependent variable. Test scores are standardized with mean zero and standard deviation of one. Column (1) are the main results, column (2) native students, column (3) first and second generation immigrants, column (4) students with low educated parents (¡= 12 years), and column (5) students with higher educated parents (¿= 14 years). All regressions include subject fixed effects and a control for teachers' subject-specific specialization degree in either math or science. PISA sampling weights are utilized. Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

native to their country of residence (column (2)), it is the first-mentioned group which has the largest advantage of having teachers with on-the-job training. In fact, the effect is almost five times as large for those with foreign backgrounds. This group seems to be a driver of our estimated effect. This heterogeneity may have multiple explanations. One is that the on-the-job training specifically trains teachers to become better teachers for children with foreign backgrounds. Another potential explanation is that students with migrant backgrounds, who typically perform worse than native-born students, have more unrealised potential that is unlocked by on-the-job training. Another way to phrase it is that there are decreasing returns to better teaching as learning outcomes increase.

Moving on to the heterogeneity between students of parents with high (column (4)) and low (column (5)) education, respectively, it is the children of highly educated parents who benefit more from teachers with on-the-job training. One motivation for the differences between students with low- and high-education parents may be that the parents with high education care more about their children's education and thus take more opportunities to interact with the teachers. It may be that the improvement of teaching skills is not only beneficial inside the classroom, but also that they are beneficial in communication with parents. In that case, students of highly educated parents may have an additional take-up source of the treatment.

Between the two dimensions of comparison, the results may seem somewhat contradictory. It is the typically more disadvantaged group, students of immigrant background, and the more advantaged group, students with highly educated parents, who see the largest effects of on-the-job training. Then it is important to remember that immigrants are a heterogeneous group and many of them may be highly educated. Actually, 66 percent of students with immigrant background have highly educated parents, while the same number for native students is 78 percent. It could also be that the same channel discussed in the parental education section, interest and activity in children's education, may be present also for migrant children. Lastly, we cannot rule out that our empirical strategy has not been able to get rid of all endogeneity and that this could be a driver of the un-intuitive results.

6 Possible Mechanisms and Robustness Checks

6.1 Possible Mechanisms

Due to the rich data on teacher characteristics that PISA supplies, we are able to explore different mechanisms that might explain our main results. While we cannot make causal claims when we no longer study student outcomes, the estimated correlations might help illuminate potential channels through which our estimated effect works. We hypothesise that teachers might obtain additional skills in communicating with students' parents or guardians through on-the-job training. Awareness of parents about their child's performance in school could let parents adjust their resources and lead to potential increases in educational achievement. Therefore, we estimate the effect of on-the-job training on the hours that teachers spend per week communicating with parents. Results are reported in column (1) of Table 3. They are highly statistically significant and indicate that teachers that receive on-the-job training spend on average about ten minutes per week more communicating with their students' parents. We argue that this is a plausible effect size and could be one mechanism that drives our main results. Considering that we use data from 2022 and students tested in this year were potentially affected by school closures due to the Covid-19 pandemic in previous years, the communication channel between teachers and parents could have been particularly important for the students in our sample.

A further mechanism we consider is that teachers achieve better skills in adjusting their teaching to the needs of their students through on-the-job training. For this purpose, we use the responses of teachers of how often they tailor their teaching to students' needs. We create a dummy variable that indicates 0 if teachers adjust their teaching only for some lessons or less and 1 if they adjust their teaching for many lessons or more. In column (2) of Table 3, we observe a large and significant effect of teachers' on-the-job training on their adjustment of classes to their students needs. Teachers receiving training have an approximately 23 percentage points higher share of tailoring many or more lessons to their students needs. Our results suggest that teacher adjustment could be an important mechanism to enhance students' achievement.

Table 3: Estimated Effects of On-the-job Teacher Training on Plausible Mechanisms

	(1) Communication	(2) Tailored Teaching
On-the-job Training (1=yes)	0.172^{***}	0.227^{***}
	(0.0445)	0.0664
Teacher Controls	YES	YES
Student Fixed Effects	YES	YES
N	22,772	22,772

Notes: Results from weighted student fixed effects regressions. Column (1) shows effects on hours of communication to parents. Column (2) shows effect on binary tailored teaching indicator. All regressions include subject fixed effects and a control for teachers' subject-specific specialization degree in either math or science. PISA sampling weights are utilized. Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

6.2 Investigating Coefficient Stability

We examine the coefficient stability of our estimates by exploiting Oster (2019)'s approach on selection on unobservables. This helps us to further investigate the robustness of our coefficients to an omitted variable bias. Under the assumption that the relation between unobserved traits and the treatment variable is recoverable from the relation of observed characteristics and the treatment variable, shifts in coefficients need to be accompanied by shifts in R^2 to indicate omitted variable bias (Oster, 2019).

We follow two methods to investigate the stability of our coefficients. The first one suggests to estimate a δ which represents a proportional relation between the selection on observables and unobservables. A $\delta > 1$ implies that the selection on unobservables would need to be disproportionally high compared to observables to explain away the effects of teachers' on-the-job training on students' achievement. For this reason, a $\delta > 1$ is the established critical value to identify robustness (Berniell & Bietenbeck, 2020).

The second method suggests to estimate the so-called identified set $\Delta = [\tilde{\beta}, \beta^*(\min\{R_{max}, 1\})]$. Hereby, the lower bound of the identified set, $\tilde{\beta}$, represents the coefficient the students being taught by a teacher who participated in any on-the-job training, estimated in a regression with the full set of teacher- and class-level controls. For the upper bound, the unrestricted bias-adjusted estimator for $\delta = 1(\beta^*)$ is selected. If the identified set includes zero, the coefficient stability fails the robustness check due to a remarkable change in sign of the unrestricted estimator when compared to $\tilde{\beta}$ (Oster, 2019).

Table 4: Selection on Observables: Investigating the Coefficient Stability

	(1) Uncontrolled Effect	(2) Controlled Effect
On-the-job Training (1=yes)	0.0131**	0.0128**
	(0.0059)	(0.0060)
Teacher Controls	NO	YES
N	80,648	80,648
R^2 (within)	0.008	0.008
δ for $\beta = 0$		0.671
Identified set $\Delta = [\tilde{\beta}, \beta^*(\min\{R_{max}, 1\})]$		[-0.005; 0.031]
Does the identified set exclude zero?		NO

Notes: Results from weighted student fixed effects regressions with student standardized test scores as dependent variable. Test scores are standardized with mean zero and standard deviation of one. Robust standard errors clustered at the class level are in parentheses. All regressions include subject fixed effects and a control for teachers' subject-specific specialization degree in either math or science. PISA sampling weights are utilized. Standard errors are in parenthesis (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$).

Independent of this method, we require an arbitrary chosen R_{max}, which represents the R^2 of a regression that includes all observed and unobserved controlls[3]. In practice, it is established to use a $R_{max} = 1.6\tilde{R}$, where $\tilde{R}$ is the R^2 of a regrerression with the full set of controlls (Oster, 2019).

[3]As we follow a within-student fixed effects approach, we utilize the within-R^2 rather than the overall-R^2 of the specification with all controll variables included.

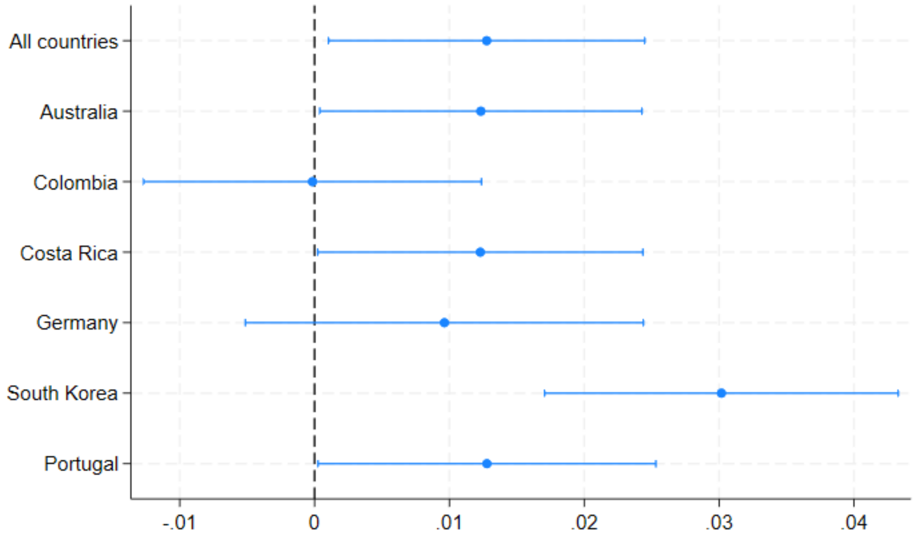

Figure 3: Estimated effect, leaving one country out at a time. 95%-CI.

Table 4 presents the uncontrolled (Column (1)) and controlled effect (Column (2)) of on-the-job training on students' outcomes. We follow the standard literature on economics of education and consider the within R^2 which indicates, how much variation of the test scores within students is explained by our fixed effects model. Our critical value, δ, lies with 0.671 below the critical value of 1. That implicates that the selection on unobservables would need to be about 0.67 times as large as the selection on the observed teacher characteristics in order to drive the treatment effect of on-the-job training to zero. Additionally, the identified set does not exclude zero for our specified value of R_{max} indicating considerable sign changes for the unrestricted bias-adjusted estimator. Overall, those findings suggests that we cannot fully exclude that omitted variable bias is potentially driving our results.

6.3 Leave-One-Out Analysis

In addition to the formal investigation of coefficient stability following Oster (2019), we also include estimates excluding one country at a time. This is to see whether there is a single country driving the results, or whether the effect is stable across borders. The resulting coefficients are plotted in Figure 3. As seen, excluding most countries does not affect the overall results. However, there seems to be a large effect in Colombia and a small effect in South Korea that impacts the results. One potential contributor for the results when leaving out South Korea is that South Korea has an exceptionally high degree of native-borns (98 percent, compared to pooled-sample average of 80 percent), which we saw in the heterogeneity analysis have lower treatment effects than foreign-borns. In other words, there seems to be large country-by-country heterogeneity and that we thus have limited external validity outside of our analysed countries. Apart from population composition, one explaination for this heterogeneity is that the actual treatment, receiving on-the-job training, may be differently designed across countries.

7 Conclusion

In this paper, we investigate the effect of teachers' on-the-job training on students' standardised test scores in order to assess teachers' effectiveness. Hereby, we draw data from PISA (Programme for International Student Assessment). Applying a within-student fixed effect approach, we focus on the 2022 test scores in the following six OECD countries: Australia, Columbia, Costa Rica, Germany, South Korea, and Portugal. Our main results suggest that on-the-job training for teachers significantly increases students' achievement by roughly 0.013 of a standard deviation. Our analysis suggests this effect occurs, among other things, through improved communication with parents and more teaching tailored to specific students. However, the proposed mechanisms cannot be interpreted causally and could be a reflection of self-selection of teachers into on-the-job training.

From a policy perspective, our estimated effect emphasizes the importance of teachers' on-the-job training. This is especially important for first and second-generation immigrant students. One suggestion for policymakers is to focus resources on on-the-job training in areas with high levels of migration. Literature in economics of education shows that teachers tend to accumulate their pedagogical skills and knowledge mainly while already teaching, especially in the first years of their profession (Clotfelter et al., 2010; Harris & Sass, 2011). As teachers usually do not switch their careers and influence students' short- and long-term outcomes, it is crucial for them to be updated about the ongoing innovations in their field as well as new pedagogical methods. Especially the COVID-19 pandemic but also the current technological development are prominent examples of the importance of obtaining new pedagogical skills. Remote lectures and assignments are nowadays a major part of a student's education and teachers need to adjust their assessment but also teaching style. One possibility to support teachers to understand and anticipate such new developments would be to equip them with intense on-the-job training as those tend to be more beneficial for teachers' effectiveness (Harris & Sass, 2011).

However, these policy implications have to be considered with caution. Although the within-student fixed effects approach addresses most of the possible endogeneity issues, we cannot perfectly rule out all potential biases. One example of that would be unobserved teacher characteristics that might not be captured by the included controls, such as teachers' motivation. Following Oster (2019)'s approach on the selection of observables and unobservables we cannot fully exclude that some omitted variables are not driving our results. However, our main results in Table 1 remain stable after adding controls which might limit some of the concerns.

Finally, this essay cannot address which specific content of on-the-job training is important for teachers' effectiveness. This opens a research gap for future studies as it could provide valuable insights for policymakers and educational institutions.

References

Angrist, J. D., & Lavy, V. (2001). Does teacher training affect pupil learning? evidence from matched comparisons in jerusalem public schools. *Journal of labor economics*, *19*(2), 343–369.

Berniell, I., & Bietenbeck, J. (2020). The effect of working hours on health. *Economics & Human Biology*, *39*, 100901.

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

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

Bressoux, P., Kramarz, F., & Prost, C. (2009). Teachers' training, class size and studentsâ outcomes: Learning from administrative forecasting mistakes. *The Economic Journal*, *119*(536), 540–561.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American economic review*, *104*(9), 2593–2632.

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

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

Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? panel estimates from pisa. *Journal of Development Economics*, *104*, 212–232.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of public economics*, *95*(7-8), 798–812.

Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (2014). Teacher effects and teacher-related policies. *Annu. Rev. Econ.*, *6*(1), 801–825.

Jacob, B. A., & Lefgren, L. (2004). The impact of teacher training on student achievement: Quasi-experimental evidence from school reform efforts in chicago. *Journal of Human Resources*, *39*(1), 50–79.

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

Mullis, T. L., Martin Michael. (2020). 20 years of timss: International trends in mathematics and science achievement, curriculum, and instruction [Accessed: 2024-04-19]. http://therisingsea.org/notes/FoundationsForCategoryTheory.pdf

OECD. (2020). *Talis 2018 results (volume ii)*. https://doi.org/https://doi.org/https://doi.org/10.1787/19cf08df-en

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

A Appendix

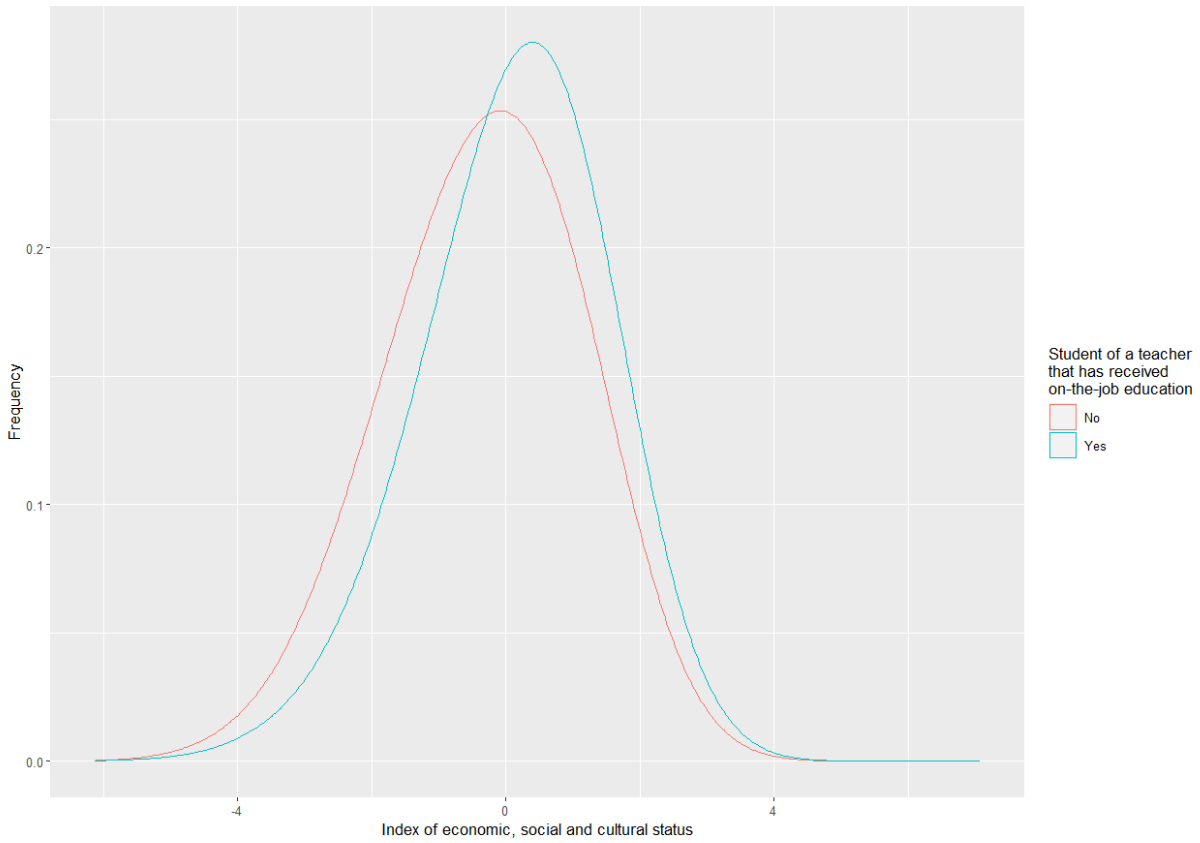

Figure 4: Frequency of students taught by teachers participating in on-the-job training programs

Table 5: Selection of countries with a clear teacher student connection rule

		Age			Experience			Upper Edu.	On the Job Edu.	Math Scores			Science Scores		
	Country	Mean	Median	SD	Mean	Median	SD	Ratio	Ratio	Mean	Median	SD	Mean	Median	SD
1	AUS	41.43	41.00	12.43	14.46	12.00	11.24	0.33	0.98	491.90	488.13	96.12	511.66	512.68	102.13
2	COL	43.81	41.00	10.02	16.71	15.00	9.36	0.58	0.49	384.10	386.20	72.66	412.88	414.81	85.16
3	CRI	38.26	39.00	8.30	13.60	14.50	7.40	0.87	0.78	381.52	376.80	62.72	409.00	402.61	75.91
4	DEU	44.11	44.00	11.16	16.03	14.00	11.01	0.79	0.63	480.87	484.95	92.05	497.70	503.99	101.53
5	KOR	46.54	45.00	8.14	19.39	18.00	9.24	0.50	0.66	525.84	534.01	102.74	525.17	537.87	97.07
6	PRT	51.48	52.00	7.46	25.10	27.00	9.36	0.25	0.56	485.81	489.58	80.86	497.40	503.59	82.30

Notes: Statistics are obtained using sample weights provided by PISA.