Organisation for Economic Co-operation and Development

ECO/WKP(2023)34

Unclassified **English - Or. English**

22 December 2023

ECONOMICS DEPARTMENT

QUANTIFYING THE EFFECT OF POLICIES TO PROMOTE EDUCATIONAL PERFORMANCE ON MACROECONOMIC PRODUCTIVITY

ECONOMICS DEPARTMENT WORKING PAPERS No. 1781

By: Balázs Égert, Christine de la Maisonneuve and David Turner

All Economics Department Working Papers are available at www.oecd.org/eco/workingpapers

JT03535295

Quantifying the effect of policies to promote educational performance on macroeconomic productivity

ABSTRACT/RESUMÉ

This paper evaluates the link between educational policies and i) student performance and ii) macroeconomic measures of productivity. The analysis has two stages. First, using the 2015 and 2018 PISA databases, it quantifies the relationship between student test scores and the characteristics of students taking the tests, their school environment and national educational systems. Second, assuming that these relationships reflect the effect of different characteristics/policies on student test performance, the second stage converts the latter into an estimated effect on macroeconomic measures of productivity using a new measure of human capital as an intermediary variable. This new measure of human capital, devised in previous OECD work, combines student test scores and mean years of schooling with estimated elasticities that suggest the former is more important. The analysis shows a positive association between spending on education and student test scores, but only for levels of student expenditure below the OECD median, suggesting scope for currently low-spending countries to raise student performance with potential gains to long-run productivity. Boosting participation in early childhood education as well as improving teacher quality is found to generate large aggregate productivity gains. There are significant, but smaller, macroeconomic gains for many countries from limiting grade repetition and ability grouping across all subjects as well as increasing the accountability of schools. Finally, the results provide evidence for income inequality having a major influence on productivity through a human capital channel.

Keywords: PISA, education policies, student test scores, productivity, human capital, OECD
JEL: E24, I20, I25, I26, I28

****

Quantifier l'effet des politiques visant à promouvoir les performances éducatives sur la productivité macroéconomique

Ce papier évalue le lien entre les politiques d'éducation et i) les performances des élèves et ii) les mesures macroéconomiques de la productivité. L'analyse comporte deux étapes. Premièrement, à l'aide des bases de données PISA 2015 et 2018, il quantifie la relation entre les résultats des élèves aux tests et les caractéristiques de ces élèves, leur environnement scolaire et les systèmes éducatifs nationaux. Deuxièmement, en supposant que ces relations reflètent l'effet de différentes caractéristiques / politiques sur les performances des élèves aux tests, la deuxième étape convertit ces dernières en un effet estimé sur les mesures macroéconomiques de la productivité en utilisant une nouvelle mesure du capital humain comme variable intermédiaire. Cette nouvelle mesure du capital humain, élaborée dans le cadre de travaux antérieurs de l'OCDE, combine les résultats des élèves aux tests et la durée moyenne de scolarité avec des élasticités estimées qui suggèrent que les résultats des élèves sont plus importants. L'analyse montre une association positive entre les dépenses d'éducation et les résultats des élèves, mais uniquement pour les niveaux de dépenses par élève inférieurs à la médiane de l'OCDE, ce qui suggère qu'il est possible pour les pays actuellement peu dépensiers d'améliorer les performances des élèves avec des gains potentiels de productivité à long terme. Stimuler la participation à l'éducation de la petite enfance et améliorer la qualité des enseignants génère d'importants gains de productivité globale. De nombreux pays pourraient obtenir des avantages macroéconomiques importants, quoique moindres, de la limitation du redoublement et du regroupement des compétences, ainsi que de l'augmentation de la responsabilité des écoles. Enfin, les résultats démontrent que les inégalités de revenu exercent une influence majeure sur la productivité via le canal du capital humain.

Mots clés: PISA, politique d'éducation, résultats du test des étudiants, productivité, capital humain, OCDE
JEL: E24, I20, I25, I26, I28

Table of contents

Tables

Figures

Boxes

Quantifying the effect of policies to promote educational performance on macroeconomic productivity

By: Balázs Égert , Christine de la Maisonneuve and David Turner [1]

> *"Our top priority was, is and always will be education, education, education" (Prime Minister Tony Blair, 2001)*

1. Introduction

1. Policymakers often emphasise the importance of education. Moreover, human capital is widely regarded as a fundamental driver in the theoretical literature on economic growth (Jones, 2016), although finding an empirically robust link between human capital and macroeconomic growth has been difficult (OECD, 2022a). This paper follows up on previous OECD research, which constructed a new measure of human capital that was found to have a link with macroeconomic productivity (Egert et al., 2022). This measure of human capital combines student test scores and mean years of schooling and suggest the 'quality' dimension of education (as captured by student test scores) is more important than the quantity dimension. In contrast to the existing literature, the relative weights of quantity and quality are not imposed or calibrated but directly estimated. This paper connects policies to the stock of human capital by using data from the OECD's 2018 Programme for International Student Assessment (PISA) to quantify the relationship between policies and student test scores. This helps quantify the effect of policies impacting educational performance on human capital and macroeconomic productivity.

2. Data on student test scores from PISA provides the main building block for the new measure of human capital. PISA is a global programme to evaluate educational systems in OECD member and non-member countries by measuring the performance of 15-year-old students in standardised tests in mathematics, science and reading. It not only collects student test scores, but also records many of the characteristics of the students taking the tests, as well as of their school environment and national educational systems. This has provided researchers with a series of rich data sets with which to quantify the relationship between such characteristics, many of which are potentially influenced by policy choices, and student educational performance. The main contributions of this paper are: firstly, that it analyses this relationship for the most recently published PISA datasets of 2015 and 2018, which have thus far been

[1] The authors are members of the Macroeconomic Analysis Division of the OECD Economics Department. They would like to thank Luiz de Mello, Alain de Serres and Zuzana Smidova in the Economics Department, Paulo Santiago, Karine Tremblay, Francesco Avvisati and other colleagues from the OECD Directorate for Education and Skills, and delegates of Working Party No. 1 of the OECD Economic Policy Committee, for useful comments and suggestions. Many thanks also to Veronica Humi for editorial assistance.

much less explored in the literature; secondly, the effect of different characteristics/policies is quantified not only in terms of their effect on student test performance, but also in terms of their effect on macroeconomic measures of productivity using the new measure of human capital as an intermediary variable.

3. The remainder of this paper is organised as follows. Section 2 describes the modelling framework, deals with data issues and definitions, and explains how the estimation results are used to quantify policy effects. Sections 3 to 7 review the literature and discuss the new estimation results for the following policy areas: early childhood education (Section 3); teaching resources in primary and secondary education (Section 4); tracking and grade repetition (Section 5); school governance issues including school autonomy and accountability (Section 6); and the role of socioeconomic background and income inequality (Section 7). Finally, Sections 8 and 9 summarise the results and provide some concluding remarks.

2. The modelling framework and dataset

4. A large body of published research has looked into the drivers of student outcomes at the national, sub-national and cross-country levels based on student test scores and associated data collected under PISA (see Hanushek and Woessmann, 2011; and Smidova, 2019, for literature reviews) , though there is a growing number of studies exploiting the more recent 2015 and 2018 PISA vintages (see Hippe et al. 2018; Murat and Bonacini, 2020; Strello et al., 2021; De Philippis and Rossi, 2021; Avvisati and Givord, 2023; and Dong and Cornachione Kula, 2023). The current paper contributes to this growing body of literature and focuses on the 2015 and 2018 vintages of PISA data, exploiting around 140,000 student-level observations across 25 OECD countries.[2] This section overviews the empirical estimation framework and discusses issues relating to empirical modelling, variables' selection and data definitions.

2.1. The modelling framework

5. To quantify the link between education policies and student test scores, this paper follows the approach of Boarini (2009), Boarini and Lüdemann (2009), and Fuchs and Woessmann (2018) and estimates an education production function, which connects educational outcomes, measured by the OECD's PISA test scores, to inputs including educational policies measured at the individual and school levels, and a range of student and school characteristics, as follows:

$$PISA_{i,s,c} = \sum_{a=1}^{b} \alpha_a \, PI_{a,i,s,c} + \sum_{d=1}^{e} \beta_d \, PS_{d,s,c} + \sum_{f=1}^{g} \gamma_f \, PC_{f,c} + \sum_{h=1}^{j} \delta_h \, INT_{h,i,s,c} + \varepsilon_{i,s,c}, \qquad (1)$$

where:

- $PISA_{\mathrm{i,s,c,sub}}$ denotes the PISA test score for individual i in school s and country c for subject matters science, maths, reading and the arithmetic average of the three.
- *PI and PS* stand for educational policies measured at the individual and school levels, but also for individual and school-level characteristics.
- *PC* stands for country-level policies.
- *INT* denotes interactions between individual-, school- and country-level policies.

6. The regressions are run for the 2018 PISA dataset, with the 2015 vintage used as a robustness check and because the variables relating to school autonomy are only available in the earlier vintage. Because regression results based on the 2015 and 2018 vintages are broadly similar, the main text shows

[2] Though the PISA dataset includes more countries and observations, some covariates considered in the empirical analysis are not available for a number of countries, thus limiting the final estimation sample size.

and interprets estimation results only for 2018 (except for school autonomy). The full set of the 2015 results and its comparison with 2018 are reported in Annex A (Table A1 to A3).

7. The estimations are carried out using OLS regressions at the student level[3]. The robustness of equation (1) is verified in two ways: first, interaction terms are dropped, and country-level policy variables are replaced by country fixed effects[4]; second, country-level variables are added back to the regressions, but interaction effects are excluded.[5]

2.2. The explanatory variables

8. The policy and non-policy variables relevant for the empirical analysis can be divided into three categories depending on whether they are measured at the student, school or country level. Nevertheless, for the policy analysis, it is more convenient to group variables by types of policies and features they represent rather than by their level of aggregation (Table 1), which is also consistent with how subsequent sections of the paper are organised.

9. A first group of variables covers early childhood education and care (ECEC). More precisely, they measure the time students taking the PISA tests had spent in early childhood education. Three categories are used: i.) students having spent one year in ECEC, ii.) students having spent two to three years in ECEC, and finally iii.) students having participated for four to five years in ECEC. These variables are dummy variables and measure the effects on individual students relative to the absence of participation in ECEC. This detail represents an improvement over earlier PISA surveys that had only one variable on ECEC, namely whether students had spent at least one year in ECEC.

10. A second group of variables refers to teaching resources used as inputs in primary and secondary education. These variables can be further differentiated along several dimensions: i.) education spending per student at the country level; ii.) teacher quality measured by teacher qualification, e.g. the share of teachers with master’s degree (at the school level); iii.) school size and average class size in school, and iv.) staff shortages as perceived by school headmasters.

11. A third group relates to other education policies, not having immediate resource implications. This includes tracking or streaming, measured, first, by ability grouping within schools, and second, at the country level, via the first age of tracking. It also includes variables summarising the prevalence of grade repetition in primary and lower secondary education.

12. A fourth group of variables relates to school governance. It includes school autonomy in terms of budgeting, personnel decisions and curriculum (for the 2015, but not for the 2018 PISA dataset), and school-level accountability (by posting results publicly). Accountability is also measured at the country level through the existence of centralised external exit exams.

[3] Student observations are weighted with their sampling probabilities. Missing data (under 5% for most variables) have been treated by imputing plausible values. In practice, when a value is missing, it is replaced either by the weighted (student final weights) school average or the weighted country average when the school average is missing. Standard errors in the regressions are computed using balanced repeated replicate (BRR) weights following the PISA data analysis manual.

[4] Results reported in Table A2 in Annex A show that regressions including country-level variables and interaction terms explain equally well the variance in the dependent variable (PISA scores) compared to regressions including country fixed effects.

[5] A prominent feature of the cross-sectional regressions set out in (1), widely acknowledged in the research community (see e.g. Woessmann, 2018), is that they cannot capture causal relationships. Hence, the interpretation of the results presented later in this paper will avoid causal interpretation.

13. A fifth group refers to social background and inequalities, including: i) the socioeconomic background of individual students, reflected in parental education, parental occupation and home possessions, including books; ii.) dispersion in the socioeconomic background within school (measured as the standard deviation of socioeconomic background of individual students in each school.); and iii.) income inequalities at the country level, measured by the Gini coefficient.

14. Finally, a set of control variables for other student characteristics such as gender, measured by a dummy variable differentiating between girls and boys at the individual level and, a dummy capturing if a student speaks a foreign language at home[6], and the age of the students at the time the PISA tests were administered.

15. These broad groups are identified based on past research. It should be noted that the PISA surveys include a large variety of variables that are often strongly correlated with each other and hence cannot be used simultaneously in regression analysis, which also has implications for interpreting results. For instance, indicators measuring the quality of teaching such as teachers' experience, attendance of training programmes, salaries and the qualities of school headmasters are strongly correlated with the share of teachers with master's or postgraduate degree.

16. In addition, several plausible interaction effects are investigated in the empirical analysis. For instance, the literature suggests that accountability might prove necessary for school autonomy to improve student outcomes. Socioeconomic background might also amplify or attenuate other policy effects. For example, interactions are considered to analyse whether students coming from better-off or from more disadvantaged families benefit most from better teachers.

[6] This variable indicates that the language spoken at home is different from the language of instruction at school. In multilingual countries, this is not necessarily a foreign language.

Table 1. Description of the variables used in the regression analysis

	Level of Aggregation	Dummy variable	Min	Max	Mean	Standard deviation*	Definition
Dependent variables							
PISA score, science	Individual		65.1	886.1	488.0	92.8	Plausible Value 1 in science.**
PISA score, math	Individual		64.3	888.1	488.2	89.5	Plausible Value 1 in mathematics.
PISA score, reading	Individual		54.7	887.7	486.5	98.6	Plausible Value 1 in reading.
Average of 3 scores	Individual		142.8	843.5	487.6	88.3	Unweighted average of 3 scores.
Independent variables							
1) Early childhood education (ECEC)							
1 year in ECEC	Individual	√	0	1			1 if one year spent in ECEC.
2 to 3 years in ECEC	Individual	√	0	1			1 if two to three years spent in ECEC.
4 to 5 years in ECEC	Individual	√	0	1			1 if four to five years spent in ECEC.
2) Resources devoted to education							
School size	School		5	8150	793	440	Number of students.
Average class size	School		13	53	27	5.5	Number of students.
Staff shortage	School		-1.9	4.0	0.1	0.9	Measuring school principals' perceptions of factors hindering teaching. Positive values mean a shortage of teaching resources relative to OECD average.***
Teachers quality							
Teachers with a master's degree	School		0	1	0.4	0.2	Share of teachers in school with a master's degree.
Education spending	Country		2319	20403	8427	3570	Total annual spending per student, average in primary (in 2011) and lower secondary school (in 2015), in PPP USD.
3) Tracking and grade repetition							
Schools with ability grouping for all subjects	School	√	0	1			1 for school with ability grouping in all subjects.
First age of tracking (in 2015)[a]	Country		10	16	14.6	2.1	First age of tracking at the country level.
Repeated once in primary	Individual	√	0	1			1 if a student repeated one grade in primary.
Repeated 2 or more in primary	Individual	√	0	1			1 if a student repeated two or more grades in primary.
Repeated once in lower secondary	Individual	√	0	1			1 if a student repeated one grade in lower secondary.
Repeated 2 or more in lower secondary	Individual	√	0	1			1 if a student repeated two or more grades in lower secondary.
4) School governance:							
Accountability							
Posting achievement data publicly	School	√	0	1			1 if school posts test results publicly, 0 otherwise.
External exit exam (in 2015)[a]	Country		0	4	2	1.4	0 if no centralised external exit exam in lower and upper secondary and 4 if centralised external exit exam in 4 or more subjects both at the lower and upper secondary levels. ****
Autonomy (in 2015)	School						
Budgeting	School		-2.9	1.1	0.0	1.0	First principal component of hiring, firing, starting salary and salary increases.
Staffing	School		-2.1	2.8	0.1	1.2	First principal component: formulating and allocating the school budget.
Curriculum	School		-3.3	1.4	0.0	1.0	First principal component: courses offered, course content, textbooks used.
Public school	School	√	0	1			1 if school is managed directly or indirectly by a public education authority, government agency, or governing board appointed by government or elected by public franchise

Table 1. Description of the variables used in the regression analysis (*cont.*)

	Level of Aggregation	Dummy variable	Min	Max	Mean	Standard deviation*	Definition
5) Socio-economic background and income inequality							
Socio-economic background (country demeaned)	Individual		-7.7	4.5	0.0	0.9	Derived as highest parental education, highest parental occupation, and home possessions including books at home. Higher values denote better family background. Normalised around a zero mean in each country.
Dispersion in socioeconomic background within schools	School		0	2.8	0.8	0.2	Standard deviation of individuals' socioeconomic background in each school.
Income inequality (in 2015)[b]	Country		25.4	51.0	34.4	6.8	Gini index, 0 represents perfect equality, 100 implies perfect inequality.
GDP per capita (in 2015)[c]	Country		11.8	107.4	39.0	17.3	In 1000s 2017 PPP USD
Gender	Individual	√	0	1			1 for girls, 0 for boys.
6) Miscellaneous							
Age	Individual		15.1	16.3	15.8	0.3	Age of students at time of PISA tests.
Foreign language spoken at home	Individual	√	0	1			1 if the language spoken at home is different from the language of the test.
Teachers picking on students (in 2015)	Individual		1	4	1.4	0.9	Student perception of whether teachers picking upon him/her. *****

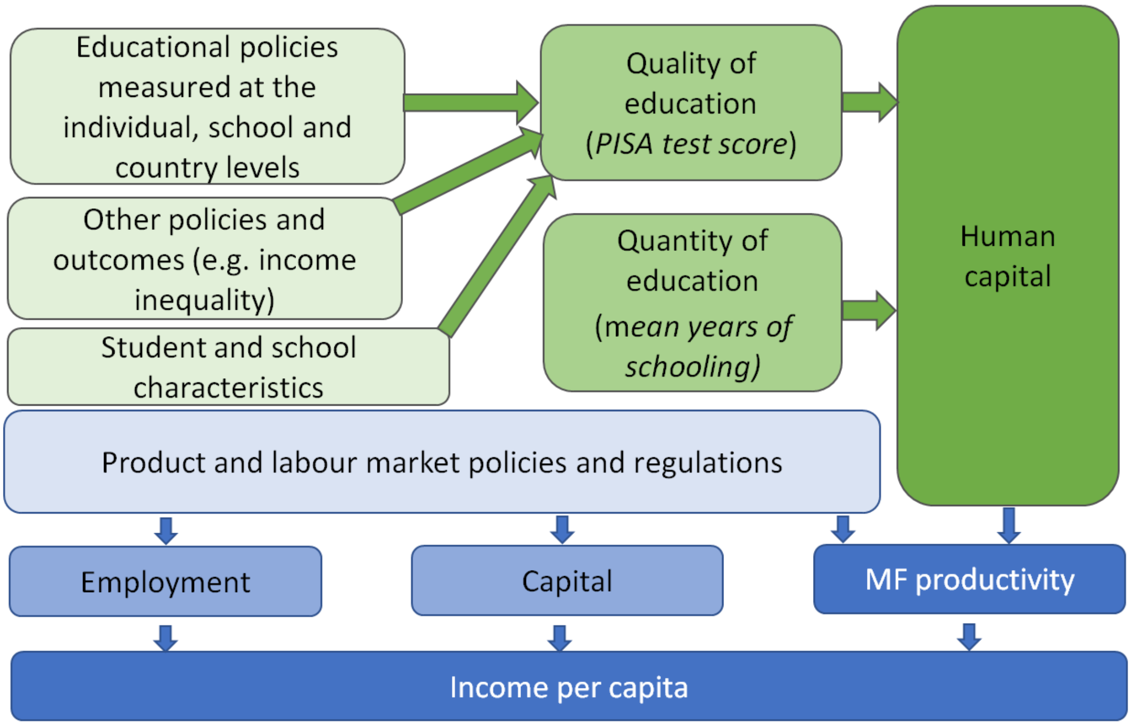
Note:
*Standard deviations of variables at the individual or school levels are calculated as the mean of within-country standard deviations. Standard deviations of variables at the country level are calculated as standard deviations of country-level variables.
**Plausible values are a selection of likely proficiencies for students that attained each score. For each scale and subscale, ten plausible values per student are included in the database. For further information on how plausible values are calculated, see the PISA Data Analysis Manuel: SPSS, Second Edition (https://doi.org/10.1787/9789264056275-en) .The first plausible value is chosen as the 10 different plausible values were giving essentially the same results in the regression analysis.
***The question asked is: "Is your school's capacity to provide instruction hindered by any of the following issues? a lack of teaching staff; inadequate or poorly qualified teaching staff; a lack of assisting staff; inadequate or poorly qualified assisting staff". The four response categories were: "not at all", "very little", "to some extent", "a lot".
****External exit exam can take the values of 0, 1, 2, 3 and 4: = 0 if no external exit exam; =1 if less than 4 subjects in external exit exams either at the lower or upper secondary level; =2 if 4 or more subjects in external exit exams at the lower OR upper secondary level or if less than 4 subjects in external exit exams either at the lower AND upper secondary level; =3 if less than 4 subjects in external exit exams at the lower secondary level AND 4 or more subjects in external exit exams at the upper secondary level; OR if 4 or more subjects in external exit exams at the lower secondary level AND less than 4 subjects in external exit exams at the upper secondary level lower; =4 if 4 or more subjects in external exit exams in lower and upper secondary.
*****Ranges from 1 to 4 where: 1 = Never or almost never; 2 = A few times a year; 3 = A few times a month; 4= Once a week or more.
Source: OECD(2018a), OECD(2018c) and OECD(2018d) except for: a) OECD (2023c); b) World Bank World Development Indicators' database and c) OECD Economic Outlook Database.

2.3. Quantifying effects at both the individual and macroeconomic level

17. The estimations allow the evaluation of the correlation between policy variables or characteristics and an individual student's PISA score, which can then be aggregated at the country-level . The effect of policies at the country level can then be related to the stock of human capital and finally aggregate productivity (Figure 1), [7] through a system that can be represented in a stylised form as equations (1) to (4).

[7] Policies might not only improve productivity but also have positive effects on well-being and health outcomes, with the latter boosting productivity.

Figure 1. The link between education policies and productivity



Note: The chart shows how educational policies are linked to PISA test score and human capital and, subsequently, to productivity and per capita income.

18. The magnitude of policy changes on individual students' test scores can be computed from the estimated coefficients on the estimated equation to explain individual PISA scores, previously described in equation (1). For assessing the magnitude of effects at the student level, it may be helpful to note that, for the 2018 vintage of PISA and the sample of countries considered here, the typical (i.e., country average) standard deviation of PISA scores across any one country is about 88.

19. Individual student or school-level PISA effects, $PISA_i$, from policy innovations are then transformed into country-level PISA scores, $PISA_c$ (equation 2). While any particular characteristic or policy may be associated with a relatively 'large' effect on an individual student's PISA score, it does not necessarily imply the potential for a large effect on a country's overall PISA score and so potential to significantly raise aggregate productivity. Rather it is first necessary to evaluate the scope for scaling up any effect across the entire population of current and future students. A particular characteristic or policy may be associated with a large effect on an individual's PISA score and be suggestive of policy improvements. However, if the share of all students with this characteristic, and/or the likely coverage of any policy change, is small, then the scope for raising aggregate productivity will be correspondingly limited. For example, grade repetition may be associated with a large effect on an individual student's PISA score, but if few students repeat a grade, then there may be little scope to exploit such an effect to raise aggregate student performance or aggregate productivity. For assessing the magnitude of effects at the country level, it may be helpful to note that, for the 2018 vintage of PISA and the sample of countries considered here, the standard deviation of average country PISA scores is about 29.

$$PISA_C = \text{Average}\ (PISA_i) \qquad (2)$$

20. Next, country-level changes in PISA student scores due to policies are translated into changes in the stock measure of human capital (equation 3). The new stock measure of human capital (HCAP) is a

cohort-weighted average of past PISA scores (denoted PISA*, where the * denotes the measure is cohort-weighted), representing the quality of education of the working age population, and the corresponding mean years of schooling, representing the quantity of education (Box 1). This cohort approach means that policy changes will only have their full long-run effect on the stock of human capital over a period of decades once the policy change has had time to impact on all cohorts in the labour force. Estimations documented in Egert et al., (2022) suggest that a sustained increase in nationwide PISA performance by one standard deviation of average country scores (about 29 PISA points) implies a long-run increase in aggregate productivity of 3.8%, or equivalently, that a sustained increase in country-wide PISA performance of about 8 points is associated with an increase in long-run productivity by 1%.

$$\ln(HCAP_C) = \alpha 1 \ln(PISA^*_C) + \alpha 2 \ln(MYS^*_C) \qquad (3)$$

21. The final step establishes changes in country level multifactor productivity (PDY) using an estimated relationship between human capital and aggregate productivity, which also includes other policy variables (X1, X2, .. etc.) that plausibly impact on productivity and have been used in previous OECD empirical work to explain productivity (equation 4).

$$\ln(PDY_C) = \beta 1 \ln(HCAP_C) + f(X1, X2, ..) \qquad (4)$$

Box 1. A new measure of human capital

This paper exploits a new measure of human capital, derived in Égert et al. (2022), that makes use of mean years of schooling and OECD data from the Programme for International Student Assessment (PISA) and the Programme for the International Assessment of Adult Competencies (PIAAC). The new measure is a cohort-weighted average of past PISA scores (representing the quality of education) of the working age population and the corresponding mean years of schooling (representing the quantity of education). Weights for PISA and mean years of schooling are estimated from the data. The idea underlying the new measure can be summarised as follows:

- PIAAC adult test scores could be used to calculate a cohort weighted stock measure of human capital. Nevertheless, PIAAC has limited country coverage and the PIAAC-based human capital measure has one observation in time, hence making it ill-suited for cross-country time series regression analysis to establish a link with productivity.
- For this reason, PIAAC adult test scores are matched with mean years of schooling and PISA student test scores of the corresponding cohort who took the student tests as 15-year-olds. PIAAC test scores are then regressed on matched PISA test scores and mean years of schooling. This approach has two important advantages. First, the estimated human capital measure covers a wider set of countries and many more years than is available for PIAAC. Second, the relative weights of the quality and quantity components are not imposed or calibrated, unlike in much of the existing literature, but are estimated directly.

Recent research conducted at the World Bank led to the development of a measure of human capital, referred to as Learning-Adjusted Years of Schooling (LAYS), that combines the quantity and quality of education (Filmer et al., 2020). The basic variant of this measure is obtained as cohort-specific mean years of schooling multiplied by relative cohort-specific student test scores (TIMSS and PISA scores) where a country's cohort's test score is compared to a high performer country (Singapore or an average of top 5 countries). Consequently, the elasticities on quantity and quality are imposed by construction to be equal rather than estimated, representing an important difference between the OECD and World Bank measures. For countries for which both LAYS, and the OECD's new measure of human capital are available, the cross-country correlation (no time series available for LAYS) between the two measures is around 0.9.

3. Early childhood education

22. Policy support for early childhood education and care[8] (ECEC) has often been viewed as a means to promote female labour force participation (Jaumotte, 2003; Thévenon, 2013), which apart from boosting labour supply can promote gender equity and poverty reduction, especially for single parents. However, another reason to support early childhood education is the effect on children's cognitive, socio-emotional, and physical development. It helps children blossom in terms of thinking, motor skills, feeling and relating to others. It also prepares them for school and life by developing social and emotional skills and some academic skills such as early numeracy and literacy. ECEC appears to matter the most for children from disadvantaged backgrounds (OECD, 2021) (see Box 2 for a review of the empirical literature).

Box 2. Early childhood education – literature review

Early childhood education has received significant attention in OECD countries over the past decades, largely initiated by Nobel-Prize winner James Heckman's work. According to this body of work, quality early childhood education (from birth to age five) of disadvantaged children has profound implications for later life by boosting the effectiveness of later education (Heckman et. al., 2013, Braga et al., 2013; Attanasio, 2015; OECD, 2022c). However, early intervention alone is not enough; it must be followed by continued investment (Heckman, 2008). Pre-school and formal education systems develop both cognitive and non-cognitive skills, which are crucial for success in life. Therefore, increasing access to and the quality of early childhood education can improve both the quantity and quality of educational outcomes, increase earnings, improve health and reduce crime. Early childhood is a critical period for brain development making investment in early childhood education more cost effective than in later life, especially if investment in education has diminishing returns with age (Cunha et al., 2006).

Heckman's research has been met with criticism on the grounds that it is predominantly based on the Abecedarian and the Perry Pre-school experiments carried out in the United States for disadvantaged children aged three to four over a period of two years, and so cannot be generalised because of the small scale of the experiments and because randomisation had not been carried out carefully enough. The curriculum in these experiments also differs from the education programmes in modern pre-schools (Weiland et al., 2023). More recent and general datasets, relying on fully randomised controls, seem to show that early childhood education effects might be smaller and shorter lived (Whitehurst, 2017; VerBruggen, 2019). This new evidence suggests that early childhood education is important, but it has to be minimally good and later stages of education are equally important. It also suggests that early childhood education is very important for the extremely disadvantaged, and those not speaking the official language but less so for children from other families (Phillips et al., 2017).

Country-specific international evidence is mixed. Starting pre-primary education early is reported to raise student test scores in Spain and Greece (Felfe et al., 2015), and boosting access to nursery schools and kindergartens in Japan was associated with improved education outcomes in later life (Akabayashi and Tanaka, 2013). Expanding universal kindergarten in the United States is found to

[8] Early childhood education and care can be divided into two parts: i.) early childhood educational development for children aged zero to two; and ii.) pre-primary education for children aged three years until they start primary school. The rest of the paper will use early childhood education and care (ECEC) for both parts.

reduce repetition rates in schools (Cascio, 2009). In France, one-year old children, especially those from disadvantaged families, attending 'crèche' are reported to perform better in language skills, though effects were mitigated for motor skills and behaviour at the age of two (Berger, 2021). Other research suggests that one additional year of pre-primary school had no significant effect on later educational achievement in France (Goux and Maurin, 2008) Similar effects are reported for New Zealand and Quebec/Canada (Hodgen, 2007; Baker et al., 2008).

Much of the literature looking at the OECD's cross-county PISA results is more encouraging. For instance, students who attended pre-primary education for more than one year scored better in the 2003 PISA survey than their peers who attended for less than a year. Using recent PISA surveys that provide more granular data on pre-primary education, OECD research (OECD, 2020) indicates that attending pre-primary education for one, two and three years improved PISA scores substantially.

23. Recent literature stresses the importance of the quality of ECEC reflected in factors such as staff-child interactions, educational activities, working conditions or staff remuneration matters considerably, though staff-to-children ratio, staff qualifications and monitoring are found to impact less on educational outcomes (Smidova, 2019, Slot, 2018 , Bigras & Lemay, 2012). For instance, children in Denmark and the United Kingdom having participated in high-quality ECEC did better in student assessments compared to their counterparts having received lower-quality ECEC (OECD, 2022c).

Against this background, many international organisations have been advocating the importance of early childhood education. UNICEF (2023) identifies pre-primary education as a cornerstone of a child's development and offering lifelong benefits: already one year spent in pre-primary education is likely to boost future educational outcomes and reduce the likelihood of grade repetitions and drop-outs.

24. Students taking the PISA test provide information on how long they had attended early childhood education. Therefore, such data provides a snapshot relating to their experience more than ten years prior to taking PISA tests, namely when students aged 15 at the time of the test, are zero to six years old. Consequently, the 2018 PISA survey reflects the situation in 2008 and earlier. While they are a valuable source of information that will be used in regression analysis to assess how participation in early childhood education correlates with student test scores at the age of 15, they are less informative in judging the scope to change ECEC participation currently. The OECD's *Education at a Glance* provides a more recent snapshot on aggregate ECEC participation to underpin current policy advice. Despite coming from different sources, there is broad consistency between PISA survey data and official data on ECEC participation, once allowance is made for the time lag implicit in the PISA data, which provides some confidence in using coefficients estimated on PISA data in the design of policy simulations based on more recent ECEC participation data.[9]

25. Many countries have recorded substantial progress in participation rates over the last ten years or so (Figure 2). Data suggest that for a majority of OECD countries, the enrolment rate of children aged three to 5 reached high levels of above 80% in 2021. In only a few were the participation rates lower than 50%

[9] In the 2018 PISA dataset, each student is assigned one of six dummy variables depending on the number of years (one to six) spent in ECE (representing their participation around 2008 or earlier). A proxy country aggregate participation rate can be computed by weighting together these dummy variables. Thus, a proxy ECE participation rate for three-to-five-year-old children (in about 2008) is computed as a sum of: the share of students having spent one year in ECE with a weight of one-third; the share of students having spent two years in ECE with a weight of two-thirds; and the shares of children having spent three to six years in ECE with a weight of 1. This proxy aggregate participation measure derived from PISA data can be compared with official participation data sourced from Education at a Glance (EaG) for a similar historical year (in this case 2008). The cross-country correlation coefficient between the PISA proxy measure and the official data from EaG is 0.77, which confirms that the PISA sampling is broadly representative and provides greater confidence in combining regression coefficients that use the PISA dummies with current data on ECE participation from EaG for the purpose of constructing policy scenarios.

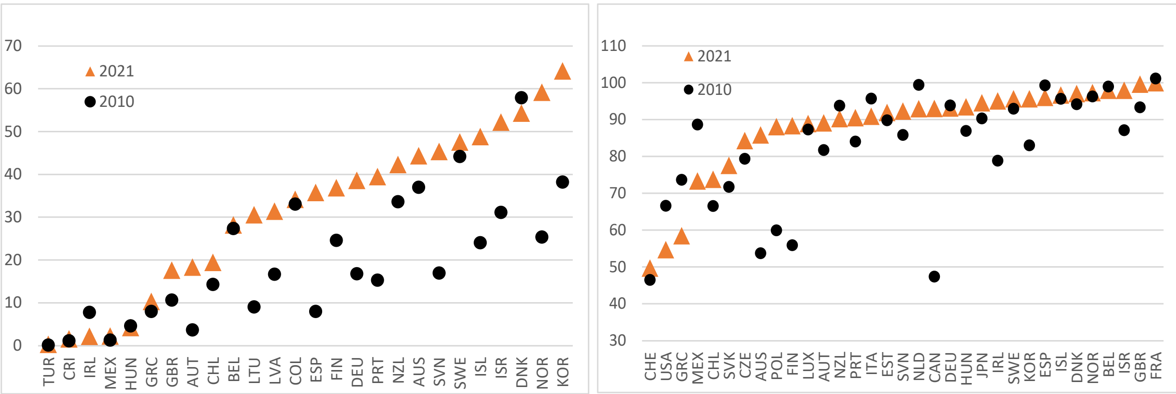
(Türkiye and Switzerland). Cross-country variation in ECEC participation of younger children aged 0 to 2 was greater. In top ranking countries (Korea, Norway, Denmark and Israel), participation rates were above 50% in 2021, whereas the average and median OECD country was at slightly above 30%. In some countries such as Türkiye, Costa Rica and Ireland, only a small fraction of young children attended early childhood education facilities (Panel A of Figure 2).

26. Whereas ECEC is instrumental in mitigating the initial disadvantage children from less well-off families face, evidence from 2018 suggests that children from disadvantaged families attended ECEC less compared to children from better-off families: their participation rate in ECEC for at least two years was 12 percentage points lower across the OECD. Data also suggest that spending on ECEC tend to be lower in comparison with spending on primary or secondary education, whilst the share of private spending is higher, hinting at some scope for boosting public investment in ECEC (OECD, 2022c).

Figure 2. Early childhood education and care in OECD countries

Enrolment rate (per cent)

Panel A. Ages 0 to 2 — Panel B. Ages 3 to 5



Source: Calculations based on OECD Education at a Glance.

Note: The first and latest data might differ slightly from 2010 and 2021. For the age group of 0 to 2, Belgium and Greece, 2019 and Portugal 2018. First year for the age group of 0 to 2, Australia, Chile, Denmark, Korea, and Sweden, 2010, Costa Rica, Greece, Israel, and United Kingdom, 2013, Türkiye, 2014, Colombia and Hungary, 2015, Belgium and Ireland, 2017. The group 3-5 includes children in primary education in countries such as Australia, Ireland, New Zealand and United Kingdom where the starting age of primary education is 5. Data for 2021 might be impacted by the COIVD-19 crisis. For some countries such as Chile, Ireland, Israel, Italy, Mexico, Spain, Türkiye or the United States, ECEC participation rates decreased compared to the pre-pandemic level.

27. More time spent in early childhood education is estimated to go in tandem with better student performance at the age of 15 (Table 2).[10] Students who spent more than one year in pre-school education achieved a higher PISA by around 34 points (a change of about one-third of an average within-country standard deviation). This result is higher than other findings in the literature. For instance, Schuetz (2009) and Woessmann (2008) find that attending pre-primary education for more than one year boosted PISA scores by between 6 to 8 points in the 2003 PISA survey. Balladres and Kankaras (2020) report that attending pre-primary education for one, two and three years boosted PISA scores by 15, 22 and 19 points.

28. While the estimated correlation is slightly more pronounced for children having spent two to three years compared to one year and four to five years in pre-school education, formal statistical testing

[10] The regression results reported here and in sections 4 to 7 are taken from regressions including all (same) policies simultaneously, reported in Table A1 in Annex A.

indicates that they are not materially different from one another. In other words, these results do not find any additional gap, in terms of higher PISA scores, from more than three years in ECEC that would relate to the attendance of zero-to-two-year-olds. Some caution is needed here because of the relatively small sample size (it relates to ECEC attendance in the early 2000s) and because it does not take into account the quality of care which is likely to be important. As shown in Box 2, some research report positive effects of expanding ECEC to children under the age of three. Moreover, there may still be an economic benefit from providing ECEC for zero-to-two-year-olds in terms of boosting female labour force participation and promoting gender equality (Jaumotte, 2003; Thévenon, 2013). Overall, these results are largely in line with the extant literature summarised earlier.[11]

Table 2. Estimation results for early childhood education

PISA=f(early childhood education, other policies, controls)

	Dependent variable: PISA score				Minimum	Maximum
	Average	Science	Math	Reading		
Independent variables						
Time spent in early childhood education and care						
1 year	33.991**	36.649**	31.518**	33.807**	0	1
2 to 3 years	34.035**	33.821**	34.363**	33.923**	0	1
4 to 5 years	25.464**	25.836**	25.203**	25.353**	0	1

Note: * and ** denote statistical significance at the 10% and 5% level. The coefficient estimates reported here are extracted from the multivariate regressions comprising a very large number of policy and control variables, reported in Table A1 in Annex A.

29. Assuming that the estimated correlation corresponds to the true causal effect, scaling it up and linking it to an improvement to human capital, suggests that a 10-percentage point improvement in the coverage of three-to-five-year-olds in ECEC will eventually raise country-level educational outcomes by the equivalent of 3 PISA points consistent with an increase in aggregate productivity by 0.4%, although it will take several decades before the full macroeconomic effect is achieved.[12] Recent data (relating to 2021) show that nearly half of OECD countries have already achieved an impressive 90% coverage for three-to-five-year-olds in ECEC. However, the same data suggest that nearly one-quarter of OECD countries have a coverage which is between 20 and 60 percentage points below this threshold, suggesting the potential for long-run gains in productivity of between 0.9% and 2.8%. Türkiye and Switzerland have the most to gain from raising ECEC participation to this 90% threshold, with potential long-run productivity gains of 2.8% and

[11] Sensitivity analysis, reported in Annex A suggests that the baseline regression results are robust to alternative specifications. Table A2 shows regressions excluding country-level variables and interaction terms. Table A3 provides two types of robustness checks. First, some variables are dropped to maximise country coverage. Second, the same specification is estimated for 2015 and 2018 (the autonomy variables, available only for 2015, are dropped). Table A4 reproduces the student-level regression results at the school level (student-level variables are aggregated at the school level). Results show that the level of aggregation (student vs. school level) does not influence the main conclusions.

[12] Back-of-the envelope calculations suggest that a 10 percentage point increase in ECEC participation of the three-to-five-year-olds would increase the employment rate by about 0.2 percentage points and per capita income level by 0.3 percent. The calculations involve two steps. First, a cross-sectional regression linking in-kind benefits (as a % of GDP) to ECEC participation suggests that a percentage point rise in ECEC goes in tandem with a 0.0117 percent increase in in-kind benefits (% of GDP). Second, using Tables 3 and 4 reported in Égert and Gal (2016) help figure out the change in employment rate and per capita income.

1.8%, respectively, but United States, Ireland, Greece, and Costa Rica might also benefit with long-run productivity gains of about 1.5% (Table 7).

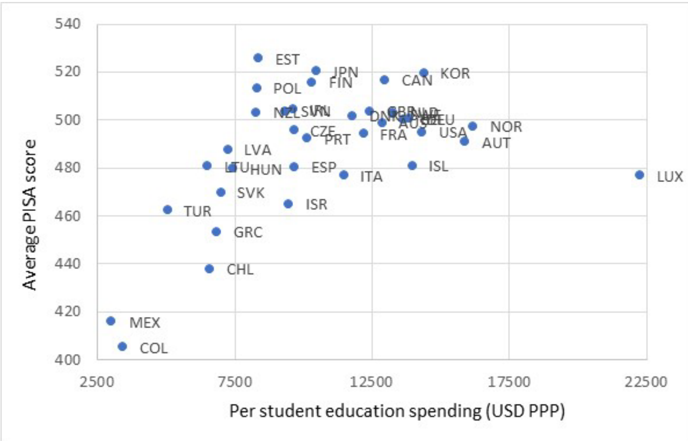
4. Resources in primary and secondary education

30. Inputs to primary and secondary education are a key driver of education outcomes. This section deals with school teaching resources, including country-level spending on education, as well as PISA survey-based data on teacher quality, staff shortages, class and school size.

4.1. Spending on education

31. The headline indicator of primary and secondary education spending per student provides a good first approximation with regard to the resources devoted to schooling. Nevertheless, the efficiency with which funds are used is of utmost importance. The literature summarised in Smidova (2019) suggests the relationship between the level of spending per student and student performance is fragile and non-linear, with any positive relationship weakening at higher levels of spending (Figure 3). For instance, education spending per capita is much lower in many OECD countries doing particularly well in student outcomes when compared to the United States (Hanushek and Luque, 2003; Woessmann, 2016). Glover and Levacic (2020) emphasise that spending per student in the Netherlands and Korea, both doing well in terms of international student test scores, was lower than the OECD average. OECD (2022b) highlights that at higher levels of education spending, the governance of funding and schools (e.g. autonomy and accountability) and its distribution are essential to ensure that funding is directed where it is most needed and most likely to lead to better teaching and learning.

Figure 3. Education spending and student outcome in OECD countries, 2018



Note: Education spending is calculated as a student-weighted average of spending per student in primary and lower secondary education.
Source: OECD Education at a Glance.

32. Regressions linking student test scores to per student education spending at the country level do not include, by construction, country fixed effects. As education spending does correlate with the level of economic development, regressions control for per capita income levels.[13] Threshold regressions are

[13] Table A2 in Annex A shows that regressions excluding country fixed effects but including a variety of country-level variables and interaction terms do explain the variance in PISA test scores as well as regressions including country

used to analyse possible non-linear effects around the sample median.[14] Per student education spending at the country level is estimated to positively correlate with student outcomes, but only at lower levels of education spending. The coefficient estimate suggests that up to a threshold defined as the median cross-country spending per student, every 1000-USD increase is associated with an 8-point increase in PISA scores (a 0.3 increase in the standard deviation of country-average PISA scores), which translates to a 1-percent increase in long-run productivity. However, the relationship is non-linear, in agreement with the existing literature, and plateaus out above the median of the 25-country estimation sample (7600 USD PPP per student) with a statistically significant negative, but trivial, effect (Table 3).[15]

33. On this basis, the two OECD countries with the lowest education spending, Mexico and Colombia, would be able to substantially boost educational outcomes and productivity by raising education spending per student to towards the OECD median: PISA scores would rise by up to 30 points (1 1/3 of standard deviation of country average PISA scores) and long-run aggregate productivity gains could reach 4%. Raising spending would also benefit countries such as Türkiye, Chile and Hungary, though to a more limited extent. However, for other countries, the gains to educational outcomes from raising aggregate spending are less obvious.

4.2. Teacher quality

34. It is widely recognised that teacher quality, including a solid knowledge of the subject taught and good pedagogical sense, is a fundamental determinant of educational success (Box 3).

35. Large disparities can be observed across OECD countries as regards the tertiary education degree of teachers in upper secondary education. In just under two-thirds of countries, the majority of teachers start their career with a bachelor's degree, whereas in just over one-quarter of countries most teachers start with a master's degree (Figure 4). In all OECD countries, teachers with a postgraduate degree, defined as all qualifications and degrees beyond master's degrees, capture a small share, with Portugal and France being the only countries with larger shares of nearly 10 %.

---

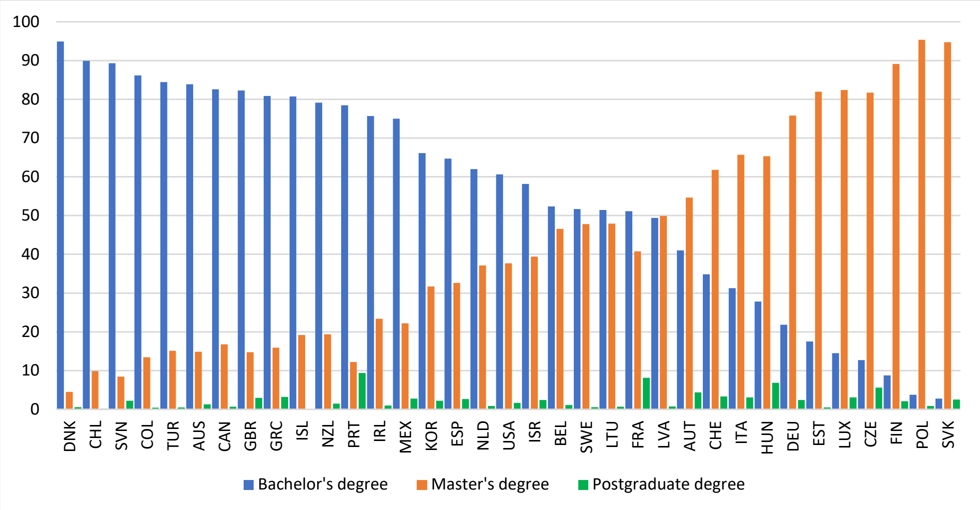
fixed effects. It is also worth noting that results do not change much whether or not the per capita income variable is included in the regressions.

[14] Estimations allow for different constants in the low- and high-education spending regimes, so as to ensure that no break occurs in the fitted data around the threshold (median).

[15] Luxembourg is excluded in alternative regressions as, according to Figure 2, this country might be an outlier. However, results for the non-linearity of education spending (and other policy effects) are unaffected by the exclusion of Luxembourg in the empirical analysis.

Figure 4. Teacher qualifications in OECD countries, upper secondary education

Share of teachers by level of diploma (per cent)



Note: Bachelor's degrees are rewarded after three to four years of studies, whereas master's degrees take two additional years to earn. Postgraduate degrees are defined as all qualifications and degrees beyond master's degrees. Data for Japan and Norway are not available.
Source: OECD Education at a Glance

Box 3. Teacher quality – selective literature review

Policies aimed at improving teacher quality should consider offering appropriate wages to teachers and a recognised status in schools and society, although the first option may be costly and the second difficult to achieve in practice (OECD, 2022c). If successfully implemented, such policies are likely to secure a larger pool of potential teachers and increase overall quality.

Better teachers will raise educational attainment as students will be more motivated to learn and study and will also help students to acquire better skills (Braga et al., 2013). School-level studies also suggest that teaching quality is one of the most important drivers of students' learning (OECD, 2005). One way to define teacher quality is by their ability to enhance student test scores. This usually goes together with solid subject knowledge, and how to teach it, positive and well-managed classes and professional values and beliefs. Empirical research found that student results are systematically better for some teachers than for others (Smidova, 2019).

Better quality teachers tend to improve long-term education outcomes and reduce student drop-out rates (OECD, 2022c). Yet measuring the quality of teaching is not straightforward. Available measures for teacher quality including years of experience, gender, qualification, education and training tend to be interrelated and often work poorly in aggregate regression analysis (Hanushek et al., 2019). Initial teacher education and initial qualifications are necessary to raise teacher quality. In addition to general teacher certification, teachers' subject-specific master's degrees in math and science, the rating of the university teachers graduate from, and their teaching experience are all correlated positively with teachers' own student test scores in secondary school. While teachers' student test scores are a good indication of skills, teachers' adult test scores, captured by the OECD's Program for the International Assessment of Adult Competencies (PIAAC) surveys, provide a more direct measure of teachers' cognitive abilities. Hanushek et al. (2019) use cross-country OECD data to show that such skills go along with better student test scores.

One important aspect of teacher quality is adequate selection of candidates in initial teacher education programmes (OECD, 2022c). In OECD countries with good PISA performance, top-performing students often become teachers and they follow lifelong learning pathways (Gomendio, 2023). These observations are corroborated by a recent literature overview by Coenen et al. (2018).

Teachers' relative salary might capture the quality of teaching, since it is well established in the empirical literature that higher wages draw more capable people to the teaching profession (Fullard, 2021). The literature also seems to indicate that higher relative wages, especially of experienced teachers, tend to be associated with better student test scores in science and math and a more positive attitude towards learning (Dolton et al., 2011; Akiba et al., 2012; Fullard, 2021). Low relative teacher salaries can lead to teacher shortages, reduce the number of qualified candidates at the entry level and increase teacher turnover, especially in schools with a higher share of disadvantaged students (OECD, 2022c). Reducing teacher shortages are, however, may reduce teacher quality, at least in the short run.

4.2.1. Estimation and simulation results

36. The share of teachers having a master's degree approximates the quality of teachers in the regression analysis. As recognised in the literature, while other measures such as experience, training, salary etc. can also characterise teacher quality, many of these variables are strongly related to one another, and therefore cannot be used simultaneously in empirical work. Consequently, results for the variable master's degree should be viewed as indicating the order of magnitude improving teacher quality might bring about and should not be interpreted literally. Bearing in mind these limitations, the estimations suggests that the maximum difference between schools, so between those with the lowest and highest number of teachers with a master's, can account for more than one-tenth of a typical within-country standard deviation of PISA scores. Transposing these numbers to the country level implies that elevating the share of teachers with master's degree from the lowest shares to the OECD median share, would be translated into productivity gains of up to 1.8% (Table 7).[16]

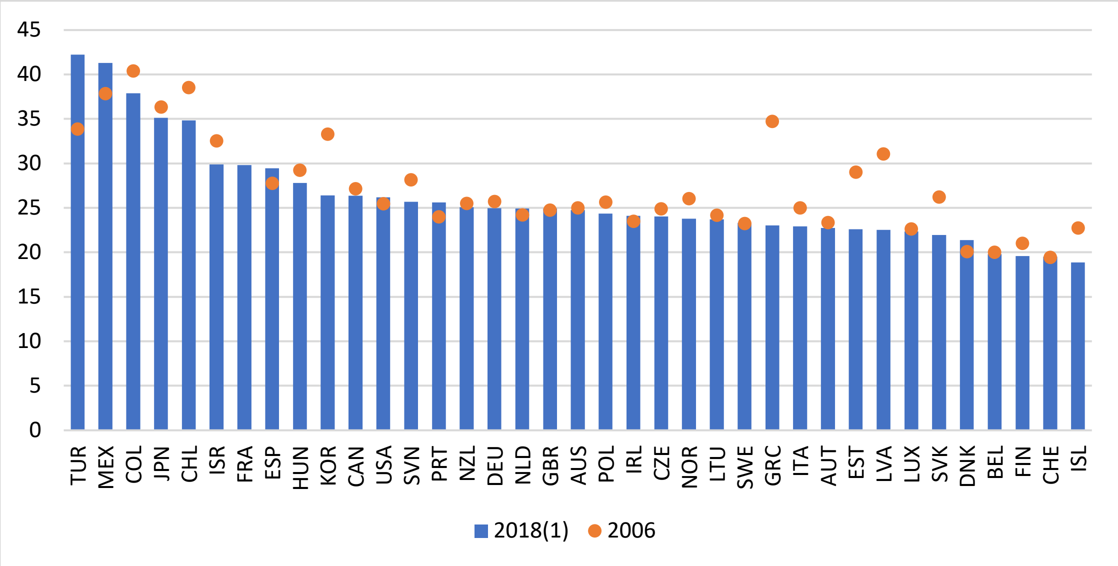
4.3. Class and school size and staff shortages

37. Although conventional wisdom suggests that smaller classes are likely to improve student performance by channelling more attention to individual students and by lowering background noise in the class, the empirical literature on the relationship between class size and PISA test scores in not very conclusive (Hanushek and Woessmann, 2011; Leuven and Ooesterbeek, 2018; OECD, 2022c); (Box 4). Around the median of 25 students per class in 2018, class size shows considerable variation across OECD countries, ranging from below 20 students per class in Belgium, Finland, Switzerland and Iceland to above 40 in Türkiye and Mexico (Figure 5). Over time, class size has decreased in two-thirds of the countries, most notably in Korea, Greece, Latvia and Estonia.

[16] Alternative catching/up scenarios could involve convergence to the top performing countries or to values above those of the top performers.

Figure 5. Average class size in OECD countries, students at the age of 15

Average number of students per class



Note: 1. Except Canada and Sweden, 2015. France has no data for 2006.
Source: *OECD Education at a Glance*.

Box 4. Class size – selective literature review

Some studies find that larger classes weigh on student test scores. A recent study by Barra and Boccia (2022) use PISA tests scores from 2000 to 2012 for OECD and non-OECD countries to show such a negative relationship. In Sweden smaller class size for students aged 10 to 13 is found to enhance student test scores at the age of 16, and education attainment and wages in adulthood (Fredriksson et al., 2013).

However, reducing class size can be more expensive relative to other policies, might entail trade-offs and the positive effects on test scores muted (OECD, 2022c). For example, a class size reduction programme in California came with a heavy bill and the positive effects were partly offset by the need to employ less well educated and trained teachers, especially in school with a larger share of students from disadvantaged families (Jepsen and Rivkin, 2009). More generally, empirical research focused on the United States suggests some evidence of positive effects of smaller classes, but the effects vary considerably across studies and tend to be not very large. The non-negligible costs implied by class size reduction are often difficult to justify by the small positive effects (Chingos, 2013). Danish data also suggest that the costs of reducing class size largely outweigh the benefits in general (Bingley et al., 2005) though the benefits are larger for disadvantaged students (Browning and Heinesen, 2007).

Others find no effect or even a positive effect from larger classes. For a sample of 649 elementary schools in the United States, Hoxby (2000) could not identify any significant relationship between class size reduction and student achievement. Woessmann (2016) reports positive coefficient estimates on class size for student-level regressions estimated for the 2003 PISA test scores for a number of OECD countries. Bigger classes correlate positively with better PISA test scores in the UK as well (Denny and Oppedisano, 2013). One explanation why larger classes might go along with better student outcomes relates to the quality of teachers. Continued professional training and lifelong learning raise the quality of teachers but also imply that teachers spend less time with students, leading to larger classes (OECD, 2022c). A different interpretation is that in highly unionised countries, teachers face less pressure to provide quality teaching and resist fiercely to deal with larger classes (Gomendio, 2023).[1] In some countries, including France and Ireland, small classes are dedicated to disadvantaged or underperforming students and schools that explains the positive link between student performance and class size, and acknowledging that students from disadvantaged families might benefit from attending smaller classes (OECD, 2022c).

----

[1] Much of the literature on how the student-teacher ratio and school size relate to student outcomes is also mixed and context-dependent. The effect may depend on other factors such as the quality of teachers, curriculum, class size and country specificities (Hanushek and Woessmann, 2011).

4.3.1. Estimation results

38. Class size turns out to have a small positive sign implying that larger classes go in tandem with improved test scores. This echoes recent findings in the extant literature. At the same time, results also suggest negative class-size effects for students with below average socioeconomic background (Table 3). Experimentation with differentiating between smaller and larger classes, in an attempt to pin down an optimal class size, was inconclusive.

39. Moving to school size, larger schools are found to go together with better test scores. This might be a source of important scale effects. It might also be because larger schools, often located in urban areas, attract better teachers and students from more advantageous family backgrounds, which would cast doubt as to whether the results could be scaled up at a national level (Table 3). Also, smaller schools are more likely to be located in rural areas where they might have multi-level classes, a challenge for teachers to manage.

40. Staff shortages, perceived by school headmasters, is estimated to be an obstacle to good educational results. A reduction in staff shortage is associated with improvements in student test scores. At the school level, considering a change from worst to best school in the sample would result in an improvement of 21 PISA points, equivalent to a 0.2 average 'within country' standard deviation.

41. Country-level simulations are not carried out for class and school size. Firstly, because it is not obvious that results can be convincingly scaled up to a national level. Consolidating schools to achieve scale effects may depend on circumstances, in particular on country size and the spatial population density of a country. Also raising class size is not a straightforward policy recommendation for many reasons including the large controversy around it in the literature.

Table 3. Estimation results for school teaching resources

Coefficient estimates on teaching resources, PISA=f(teaching resources, policies, controls)

	Average	Science	Math	Reading	Minimum	Maximum
Independent variables						
Education spending						
Below sample median	0.008**	0.006**	0.008**	0.008**	2300	7600*
Above sample median	-0.000*	-0.001**	0.001**	-0.001**	7600*	20400
Teacher qualification**	15.514**	13.770**	17.352**	15.422**	0	1
Interaction						
Teacher qualification X socioeconomic background***	11.343**	11.083**	13.168**	9.776**		
Staff shortage	-3.119**	-2.844**	-3.349**	-3.162**	-1.9	4
Class size	0.408**	0.376**	0.417**	0.432**	13	53
Interaction						
Class size X socioeconomic background***	-0.352**	-0.399**	-0.376**	-0.282**		
School size	0.009**	0.008**	0.009**	0.010**	50	81500

* 7600 corresponds to the sample median.
** Share of teachers with a master's degree.
*** Below average.
Note: As for table 2. Education spending is calculated as the average of primary and lower-secondary spending per student and is used with a lag (2011/2015 depending in data availability) given the lags with which spending in primary and lower-secondary affect educational outcomes at the age of 15.

5. Tracking and grade repetition

42. This section reviews two classes of education policies, namely tracking and grade repetition.

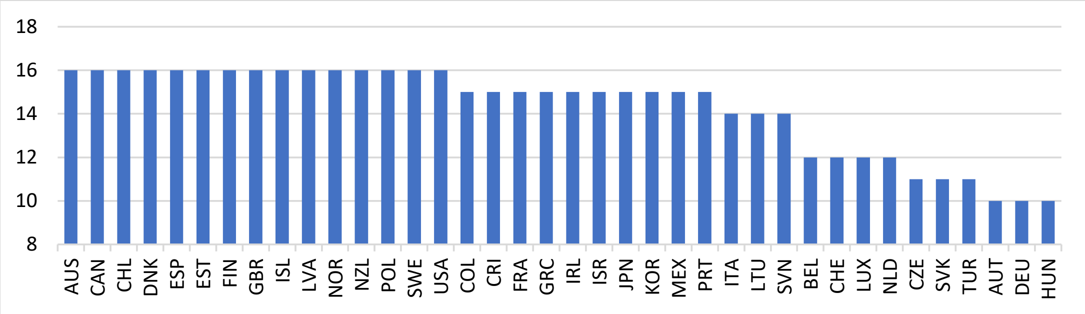
5.1. Tracking at the school and country levels

43. Tracking, the practice of dividing students by ability or achievement and by putting them into different types of schools or classes, is closely related to family background as students with lower socio-economic background are those who tend to be streamed into vocational educational tracks (OECD, 2022c), though there is no consensus about the effect of tracking on education outcomes (Box 5).

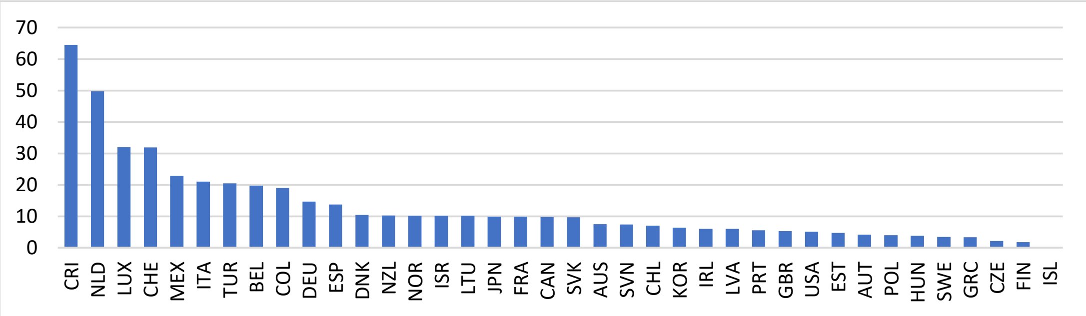
44. In a large majority of OECD countries, the first age of selection into different educational tracks occurs at the ages of 15 (10 countries) or 16 (15 countries). Only a few countries start tracking at a very early age, with Austria, Germany or Hungary being the extreme cases with as low as ten years, followed by the Czechia, Slovakia and Türkiye with 11 years and Belgium, Switzerland, Luxembourg and the Netherlands with a still staggeringly low 12 years. In terms of tracking at the school level, Costa Rica and the Netherlands stand out clearly from the crowd, given over 50 percent of schools in these countries group students in terms of ability and achievement. Numerous countries have moderate levels of school-level tracking. At the same time, in about half of the OECD countries, very few schools practice ability grouping. Iceland sets itself apart from the other countries with the absence of ability grouping, followed by the Czechia and Finland where a mere two percent of schools select and classify students based on ability and achievement (Figure 6).

Figure 6. Country- and school-level tracking, 2018

A. Tracking at country level*



B. Tracking at school level**



* First age of tracking.
** Share of schools with ability grouping for all subjects (in per cent)
Source: *OECD Education at a Glance*.

Box 5. Tracking – selective literature review

In countries where tracking occurs at an early stage, under-achieving students with poorer family background are typically channelled into specific vocational tracks at an early stage. For instance, Bonacini et al. (2021) use 2018 PISA data for 8 OECD countries with school tracking at the age of 15 or earlier (including Austria, Croatia, Germany, Hungary, Ireland, Portugal, Slovakia and Slovenia) and find that vocational schools fair worse than general schools, the differences being explained by family background. Much of the literature points out that delaying tracking until a later age helps reduce the negative effect of family background on educational attainment (Schütz et al., 2008). Though vocational education can have some advantages in the short term as it equips students with specific skills that help them secure a job soon after graduation, in the longer run, tracking can amplify initial socioeconomic inequalities, does not provide general skills including advanced cognitive and non-cognitive ones essential for re-training and lifetime learning (Liu, 2018). In other words, the initial benefits of vocational education over general schooling might evaporate over time (Hanushek et al., 2017).

Overall, empirical evidence at the country level, summarised in Smidova (2019), is mixed. On the one hand, many papers find that delaying the age of tracking improves education achievements and earnings, especially for students from disadvantaged families. At the same time, there is also substantial research showing mitigated effects or even the absence of correlation between the age of tracking and student test scores. For instance, a robust negative relationship is more apparent for reading than for math and science test scores (Hanushek and Woessmann, 2006) and adverse effects of early tracking manifest for low-achieving students but not for top students (Lavrijsn and Nicaise, 2016). A recent meta-analysis by Terrin and Triventi (2022) processing 53 papers from the last two decades confirm the no-correlation finding. The difficulty in identifying the effects of tracking might relate to the details of country-specific tracking arrangement and to the fact that only few OECD countries operate tracking starting at an early age (Figure 6).

Another and milder form of tracking sort students by ability in different subjects within schools. School-level ability separation seems to correlate with lower student test scores (Schneeweis and Winter-Ebmer, 2007). Nevertheless, whether or not grouping students by skill or achievement levels yield better results depends on the subject matter (more so for maths than for science and reading), the efficient use of time in class for low performing students and possible spillover and networking effects in mixed classes (OECD, 2018b).

45. Ability grouping at the school level shows a negative relationship with PISA scores. In other words, PISA test results tend to be worse in schools separating students by ability and achievement in all subject matters. In such schools, students scored less in the 2018 PISA tests by more than seven PISA points (corresponding to a one-tenth of a within-country standard deviation). This is much lower than the 70 PISA-point reduction for within-school ability grouping reported in Schneeweis and Winter-Ebmer (2007) for Austria.

46. Extrapolating these results to the country level by using the cross-country difference in the share of schools practicing ability grouping indicates that reducing the share of schools with ability grouping from the level of the four highest countries (Costa Rica, Netherlands, Luxembourg and Switzerland) to the OECD median would raise aggregate productivity by between 0.2% and 0.6% (Table 7).

47. Ability grouping at the school level is dragging down PISA score results of students who speak a foreign language at home instead of the official language(s) of the country. Such negative effects are sizeable: in 2018, the disadvantage for students speaking a language other than the country's official one amounts to a loss of about 15 PISA points (corresponding to one-fifth of a standard deviation in PISA scores) (Table 4). The socioeconomic status of students' family matters as well and in addition to the language spoken at home. Students who come from families with socioeconomic background below the country average and who attend schools with ability grouping tend to underperform their peers by a significant margin: more than nine PISA points (one tenth of a standard deviation).

48. Results for tracking at the country level, measured as the first age of selection, is not a robust driver of education outcomes. In 2015, delaying the first age of tracking appears to be positively related to learning outcomes. In 2018, however, pushing out the first age of selection is estimated to have no robust relationship with PISA test scores. Nevertheless, early tracking is found to be detrimental to students speaking a foreign language at home (Table 4).

5.2. Grade repetition and drop-out rates

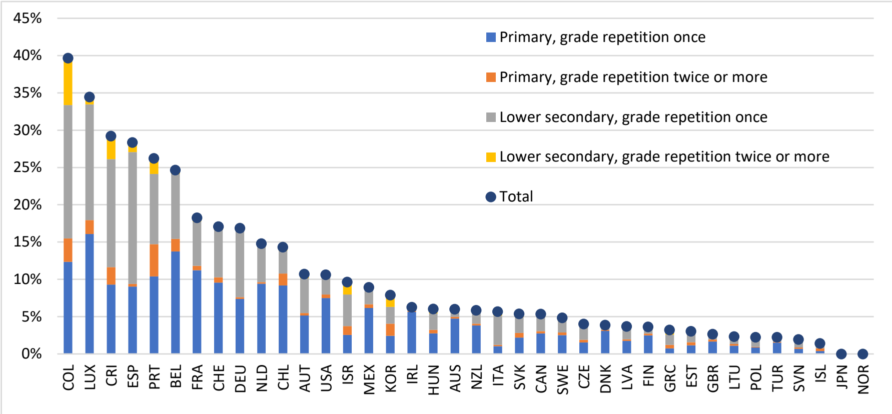
49. Grade repetition is a harsher form of sorting students by achievement and often concerns students with disadvantaged family background. The argument supporting grade retention is that it allows students an extra opportunity to acquire the knowledge necessary to move up to the next grade. Nevertheless, empirical evidence on grade repetition is overwhelmingly discouraging and improvements are short lived at the very best. At the same time, much of the literature suggests that grade repeaters underperform in PISA tests compared to students (Box 6).

50. In OECD countries with high drop-out rates such as Colombia, Luxembourg, Costa Rica and Spain (Figure 7), the macroeconomic implications might be non-negligible. A case in point is Portugal where 2009 PISA test scores in math were indeed higher for those not having repeated a grade compared to those

who had to repeat (Faria and Portea, 2016). For 2018 PISA test scores, Murat and Bonacini (2020) find that the lack of a calm place to study at home (reflecting family background-related issues) and the lack of ICT and remote work increases the likelihood of grade repetition and consequently dropping out from school in France, Germany, Italy, Spain and the United Kingdom.

Figure 7. The incidence of grade retention

Share of students having repeated a grade, 2018 (in per cent)



Source: *OECD Education at a Glance*.

Box 6. Grade repetition – selective literature review

The empirical literature suggests that children who repeat a grade underperform compared to those who did not repeat a grade (Ikeda and Garcia, 2013; OECD, 2018c). This is partly because grade repetition is not coupled with an adjusted curriculum and additional teaching resources (2022c). In addition, grade repetition lowers students' motivation and attainment expectations, increases education costs (OECD, 2023a), carries stigma, raises the likelihood of bullying victimisation (Lian, 2021) and affected students show resentment towards learning and school at the age of 15. Empirical research finds large negative effects of grade repetition on PISA scores (Hippe et al., 2018). Furthermore, grade repetition entails non-negligible fiscal costs through the additional years spent in school and a delayed entry to the labour market (OECD, 2022c). Goos et al. (2013) shows that country-level factors explain a major part in grade repetition in the (cross-country) student-level PISA data.

Grade repetition, mostly affecting disadvantaged students, substantially increases the likelihood of dropping out of school (Jacob and Lefgren, 2013; OECD, 2016; OECD, 2018c; OECD, 2022c) with adverse consequences for individuals' earnings, health and other socioeconomic outcomes in later life. For instance, in Spain, grade repetition is a good predictor of leaving school early (Gomendio, 2023). Recognising the that grade repetition is not an efficient tool to reduce gaps in educational achievements, a number of OECD countries introduced policies to limit year repetition. For instance, France initiated several programmes over the last 15 years to reduce repetition rates that had some success, though further improvement is needed (see Figure 7) (OECD 2022c). More recently, Spain, with one of the highest shares of grade repeaters in the OECD (around 30%) in lower secondary school, had implemented the 2020 Organic Law on Education that aims at making repetition exceptional, limiting the number of times a student can repeat grades and calling for targeted support to students, especially to those from disadvantaged families (OECD, 2023b). More generally, policies to mitigate dropping out of school include early detection, monitoring and prevention, already in primary school, help offered to students from disadvantaged families, and parental participation as well (Smidova, 2019).

51. Grade repetition in primary and lower secondary education is associated with poorer PISA test scores at the age of 15 (Table 4). This is very much in accordance with the existing empirical literature. For instance, students who repeat a grade once in primary school are estimated to be correlated with PISA test scores lower by about 54 points (equivalent to 2/3 of an average within-country standard deviation) This magnitude is broadly consistent with findings in the PISA literature: Hippet et al. (2018) report losses of between 30 and 70 PISA points due to grade repetition. Estimation results also suggest that the largest losses occur in primary school and that lower PISA scores due to grade retention are smaller in lower secondary. Statistical testing reveals that the effects of repeating once and twice are not significantly different from each other. Grade repeating is particularly damaging for students not speaking the national language at home, especially in lower secondary as opposed to primary school.

52. While the implications of grade repetition for individual students can be large, for the majority of countries the incidence is small enough that the significance on national educational performance is negligible. However, for a few countries, such as Colombia, Luxembourg and Costa Rica, the share of students repeating a grade is larger. Reducing the share of students who repeat one grade in primary or lower secondary of these three countries to the OECD median would raise the country average PISA test score by between 3 and 8 points (between one-tenth and one-quarter of a standard deviation), which in turn would translate into long-run productivity gains of between 0.4% to 1% (Table 7).

Table 4. Estimation results for tracking at the school and country level and grade repetition

PISA=f(tracking, grade repetition, other policies, controls)

	Average	Science	Math	Reading	Minimum	Maximum
Independent variables						
Ability grouping*	-7.005**	-8.183**	-5.688**	-7.143**	0	1
Interaction						
Ability grouping X foreign language spoken at home	-15.357**	-18.251**	-16.845**	-10.975**		
Ability grouping X socioeconomic background **	-9.298**	-9.505**	-10.221**	-8.168**		
First age of tracking***	-0.408	-0.132	-1.913**	0.820**	10	16
Interaction						
First age of tracking X foreign language spoken at home	5.321**	5.624**	4.047**	6.293**		
Grade repetition						
Once in primary school	-54.293**	-49.738**	-52.539**	-60.603**	0	1
Twice or + in primary school	-44.620**	-32.632**	-49.906**	-51.323**	0	1
Once in lower secondary school	-41.013**	-36.144**	-42.421**	-44.473**	0	1
Twice or + in lower secondary school	-24.600**	-23.497**	-26.050**	-24.253**	0	1
Interactions						
Once in lower secondary X foreign language at home	-6.989*	-10.244**	-0.053	-10.671**		
Twice or more in lower secondary X foreign language at home	-52.655**	-41.905**	-51.863**	-64.198**		

Note: * at the school level; ** below country average; ** at the country level; otherwise as for Table 2.

6. School governance issues: school autonomy and accountability

53. School autonomy and accountability are two keys, often interrelated, factors that can affect educational outcomes. School autonomy refers to the degree of freedom that schools have in making decisions about their processes, budget, personnel and curriculum. Empirical research, both based on country- and student-level data, suggests that various forms of autonomy are not strongly related to educational outcomes and provides mixed evidence with regard to the effect of school autonomy on student outcomes (Box 7). Autonomy and accountability differ largely across OECD countries (Figure 8), with country ranking across the two concepts often being quite different. Nevertheless, in some countries such as the United Kingdom, the Netherlands or the Slovak Republic, schools have a high degree of autonomy and are exposed to accountability, a combination that, according to the literature, would go in tandem with better educational outcomes.

Box 7. School autonomy and accountability – selective literature review

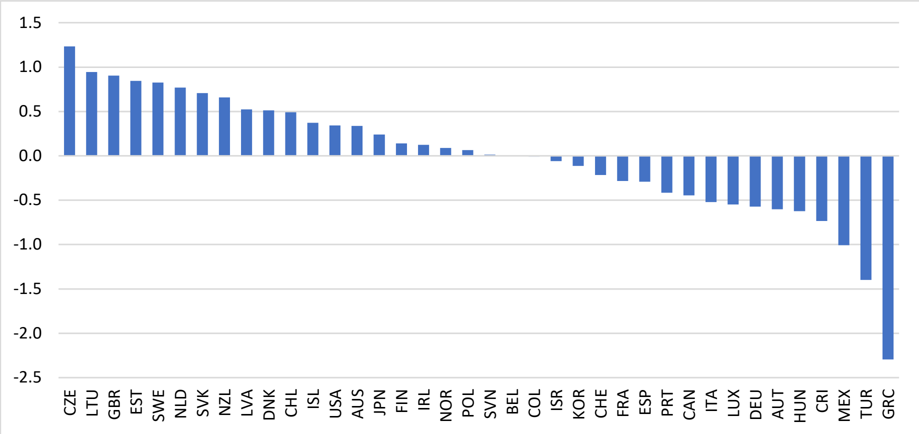
Many studies conclude that leaving more room to schools to decide on budgetary and curricular issues is often related to lower student performances (Boarini and Lüdemann, 2009; Schuetz et al., 2007; Woesmann et al., 2008; Jurges and Schneider, 2004; Hanushek et al., 2013; Hong, 2015). Evidence is more supportive of greater school responsibility for selecting teachers at the school level going along with somewhat improved PISA test scores (OECD, 2018c), though some research reports that this relationship is not a robust one either (Fuchs and Woessmann, 2007; Boarini, 2009; Boarini and Lüdemann, 2009; Hanushek et al., 2019).

Research finds that the relationship between school autonomy and student performance depends on the level of accountability (Smidova, 2019), so that schools need to be made accountable for their students' achievements and performance, which in turn incentives them to improve their teaching quality and standards. Empirical research supports the positive link between accountability and student test scores, especially when combined with autonomy. Boarini (2009) and Boarini and Lüdemann (2009) report that the publication of students' results within schools and that posting achievement data publicly are associated with better PISA test scores in sciences, though Torres (2021) cannot establish a correlation between the share of students attending schools declaring to post results publicly and PISA scores in mathematics.

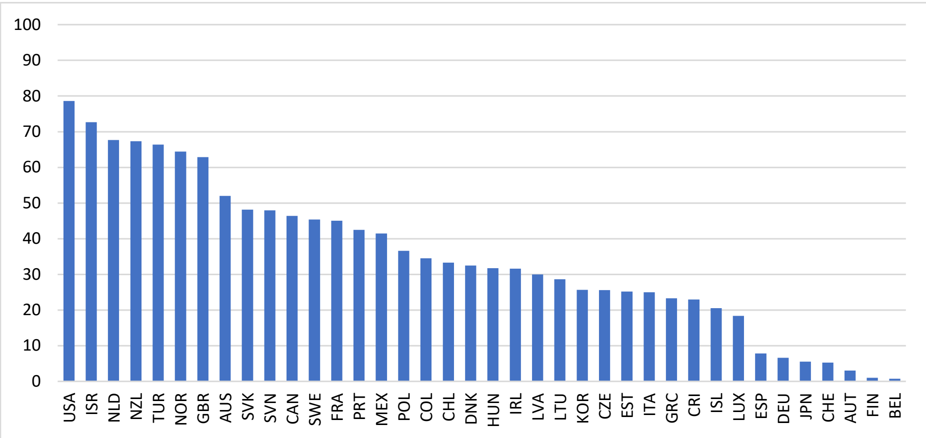
Another measure of school accountability is the use of a system of central examination, e.g. external exit exams, at the upper secondary level, that allows for an easier comparison and competition among schools for students (Braga et al., 2013; Égert et al., 2020). External exit exams are reported to correlate positively with PISA test scores (Fuchs and Woessmann, 2007; Schuetz et al., 2007, Woessmann, 2007; Woessmann 2008). Other policy levers that can increase accountability include evaluation and assessment frameworks that involve the internal and external evaluation of teachers, school leaders and schools and the reliance on a wide range of performance data (Smidova et al., 2019).

Figure 8. School autonomy and accountability in OECD countries

A. Autonomy[1]



B. Accountability[2]



1. Average level of school autonomy, 2015. Positive numbers denote more autonomy, negative numbers less autonomy.
2. Share of schools with accountable practices, 2018 (per cent).

Source: OECD Education at a Glance

6.1. Estimation and simulation results

54. Accountability (measured as schools posting achievement data publicly) goes together with significantly better PISA test scores. In 2018, individual students in schools with accountability practices are estimated to have performed better at PISA tests by about 4 points (corresponding to a 0.1 within-country standard deviation) (Table 5). Translating individual- and school-level results to the aggregate country level (via the share of schools with accountability practices) go hand in hand with relatively modest results in terms of PISA test scores and country-level productivity improvements. It is worth noting that Belgium, Finland and Austria, with a very low share of schools posting achievement results publicly, would gain the most from moving to the levels of the United States, Israel and the Netherlands.

55. At the country level, the existence of external exit exams in lower and upper secondary, thought to enhance school-level accountability, also correlates positively with PISA test scores. This correlation appears to be a non-negligible one: countries where external exit exams exist for more than three subjects

both in lower and upper secondary have average PISA test scores that are about 16 points higher[17] (about a cross-country standard deviation of country-average scores) compared to a handful of countries without such centralised external examination systems, including Belgium, Canada, Switzerland, Iceland, Japan, Mexico and Sweden. This suggests that switching to centralised external exams could lead to long-run aggregate productivity gains of 2% for this latter group of countries.

56. Three separate indicators capturing school autonomy, notably in terms of budgeting, staffing decisions and the design of taught curriculum are employed in the empirical analysis. School autonomy in budgeting and curriculum exhibits, most of the time, a positive link to PISA scores while autonomy in personal decisions appears to be negatively related to student test scores.

57. The existing body of literature suggests that autonomy might enhance student and school performance only when it is coupled with enhanced accountability. To test this finding, school-level accountability is interacted with autonomy in staffing matters. Results show some weak evidence that granting more autonomy and imposing more rigorous accountability can generate some net gains in terms of PISA test scores (Table 5).

Table 5. Estimation results for autonomy and accountability

PISA=f(school autonomy and accountability, other policies, controls)

	Dependent variable: PISA score				Minimum	Maximum
	Average	Science	Math	Reading		
Independent variables						
Accountability						
Posting achievement data publicly*						
2015	3.233**	4.814**	1.454	3.430**	0	1
2018	4.167**	4.785**	3.531**	4.184**	0	1
External exit exam**						
2015	2.587**	2.537**	1.962**	3.261**	0	4
2018	4.509**	4.121**	4.560**	4.845**	0	4
Autonomy***						
Budget	1.084**	1.196**	0.840*	1.215**	-2.9	1.1
Staffing	-2.783**	-2.471**	-3.099**	-2.779**	-2.1	2.8
Curriculum	2.601**	2.502**	2.558**	2.742**	-3.3	1.4
Interaction						
Accountablity school X autonomy staffing****	1.026*	1.881**	0.640	0.557		
Public school	-8.197**	-8.356**	-7.151**	-9.083**	0	1

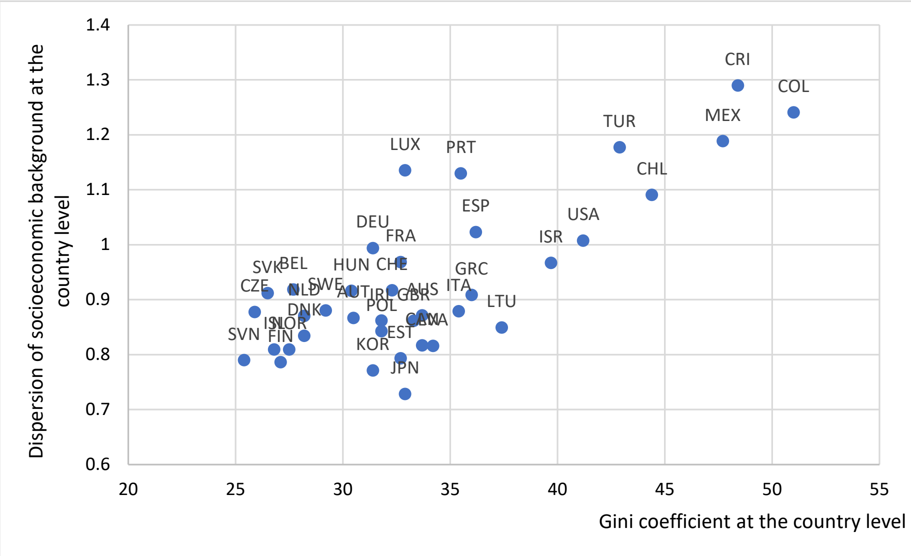
Note: * School level; ** Country level; *** School level, 2015 (data for 2018 not available); ****2015. Otherwise see Table 2.

[17] External exit exam can take the values of 0, 1, 2, 3 and 4. = 0 if no external exit exam. =1 if less than 4 subjects in external exit exams either at the lower or upper secondary level. =2 if 4 or more subjects in external exit exams at the lower OR upper secondary level or if less than 4 subjects in external exit exams either at the lower AND upper secondary level. =3 if less than 4 subjects in external exit exams at the lower secondary level AND 4 or more subjects in external exit exams at the upper secondary level; OR if 4 or more subjects in external exit exams at the lower secondary level AND less than 4 subjects in external exit exams at the upper secondary level lower. =4 if 4 or more subjects in external exit exams in lower and upper secondary.

7. Socio-economic background and income inequalities

58. The links between growth and income inequality are complex, with causation running in both directions and a range of published empirical findings (Cerra et al.,2021). However, human capital is widely accepted as an important transmission mechanism via which reduced inequality can promote higher productivity and growth (Box 8 below summaries the main findings of the recent empirical literature). Aggregate income inequality at the country level, as measured by the Gini coefficient, is positively correlated with the dispersion of socioeconomic background as defined in PISA, measured as its standard deviation across individual countries (Figure 9), so that higher levels of income inequality are mirrored in larger dispersions of students' socioeconomic background. Both measures are fairly high in Chile, Mexico, Costa Rica, Türkiye and to a lesser extent in the United States whilst both measures are comparatively low in Slovenia, Iceland, Norway, Finland and Denmark.

Figure 9. Income inequality and dispersion of socioeconomic background



Note: Vertical axis: dispersion in ESCS, horizontal axis: income inequality

Box 8. Income inequality – selective literature review

1. A number of recent studies using cross-country data (Guillemette et al, 2017; Cingano, 2014; Berg et al., 2018), all suggest a negative effect of inequality on growth, with the macroeconomic effects that are often very large.[1] Cingano (2014) also finds supporting evidence for human capital as an explicit channel through which inequality may affect growth: micro data from the Adult Skills Survey (PIAAC), suggests that increased income disparities depress skills development among individuals with poorer parental education background, both in terms of the quantity of education attained (e.g. years of schooling), and in terms of its quality (i.e. skill proficiency).

2. Further support for human capital as a transmission channel between inequality and growth, comes from the common finding in the PISA literature of a strong positive correlation between students' socio-economic background[2] and student test scores. Children from less well-off families receive less support at home for learning because of the lack of educational resources at home such as books and games, but also because parents are not willing or capable of helping in learning and offering a stimulating learning environment. Also, importantly, because parents have less financial resources to devote to individual learning activities outside of school (OECD, 2018d).

The early literature using the 2000 and 2003 vintages of PISA surveys, summarised in Hanushek and Woessmann (2011), testifies to a strong positive association between family background and student performance, although the correlation appears to differ across countries. Family background matters more in the United States and Germany but less so in Scandinavia or Canada. These findings are confirmed by more recent 2018 PISA data, with students from advantageous family backgrounds outperforming those from disadvantaged families by 89 PISA points on average (equivalent to one within-country standard deviation) (OECD, 2018d). The OECD report also underlines that countries with better overall student test scores were those where the link between socio-economic background (ESCS) and student test scores were less pronounced. Furthermore, empirical evidence also suggests that students from poorer households tend to perform worse in high income inequality countries compared to those with less income inequality (OECD, 2018d).

----

[1] For instance, estimates in Guillemette et al. (2017) suggest that a 10-point reduction in the Gini index would improve aggregate productivity by 13%. Results in Cingano (2014) imply that reducing the Gini index by 10 points would boost per capita income by 8% after five years and by 30% after 25 years.

[2] The OECD's PISA dataset collects information on students' socio-economic background, summarised in a composite indicator called economic, cultural and social status (ESCS) that captures a student's family's financial, social and human capital. It is calculated using principal component analysis applied to measures relating to i.) the highest level of parents' education, expressed in years of schooling, ii.) parents' highest occupational status based on the International Socio-Economic Index of Occupational Status (ISEI) and iii.) an index of home possessions including family wealth, cultural capital and home educational resources. The indicator is normalised to the mean of OECD countries, implying that better-off countries situate at the right of the distribution whereas less developed countries are located at the left of the distribution (relative to the OECD mean).

59. Overall, there is strong evidence suggesting that higher inequality at all levels -- individual, school and country -- go in tandem with less favourable student performance, which is very much consistent with the existing literature (Boarini, 2009, Quintano et al., 2009; West and Woessmann, 2010; Torres, 2021).

60. At the level of individual students, more favourable socioeconomic family background is found to be associated with better PISA scores (Table 6). On a scale of -8 to 4.5, a change of 1 unit in socioeconomic background goes along with an improvement of about 26 points in PISA scores (equivalent to a 0.3 average within-country standard deviation).

61. At the level of schools, larger within-school dispersion of socioeconomic background is coupled with worse student performance. A one-unit reduction in within-school dispersion would raise student test scores in the concerned schools by 25 PISA points (equivalent to a 0.3 average within-country standard deviation.

62. At the country level, a 10-point reduction in the Gini coefficient is associated with a 20-point rise in the average country PISA score (equivalent to two-thirds of a cross-country standard deviation of average country PISA scores). Based on the most recently available data on the Gini coefficient relating to 2018, if countries with the highest levels of inequality -- especially Colombia, Mexico and Costa Rica – moved towards the OECD median, the expected gains in average PISA scores would be more than one country standard deviation, which if sustained could raise long-run aggregate productivity by 4% to 5%.

63. A number of factors interact with the different measures of inequality in ways which may have important policy implications:

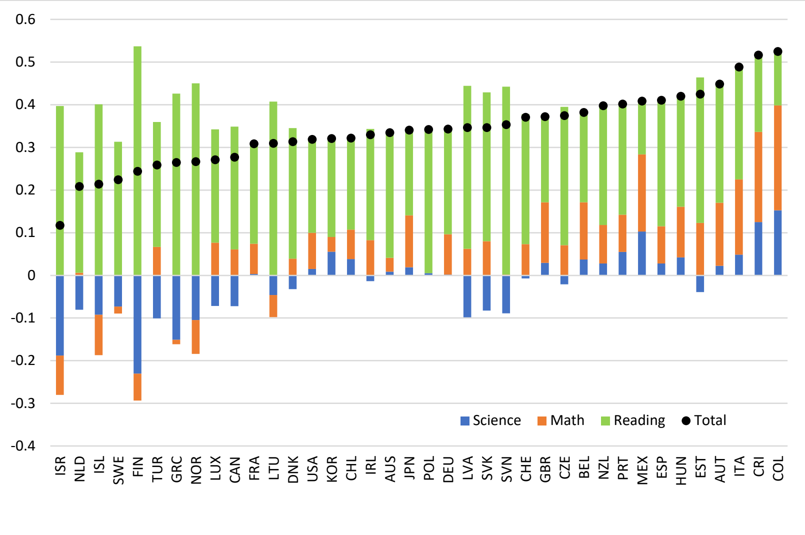
- Family background and teacher quality interact with each other: students from better-off families benefit more from better qualified teachers (with master's or postgraduate degrees) compared to students with a more disadvantageous socioeconomic background.
- Socioeconomic background interacts with country-level income inequality: in countries with more pronounced income inequalities, family background matters considerably more than in countries with moderate income inequalities. This finding is in agreement with previous OECD work, which argued that family background should have a muted effect in more equal societies (OECD, 2018e).

64. Students speaking a foreign language at home underperform substantially their peers who speak the official language with their parents and siblings. For individual students in 2018, the loss in the average PISA score amounted to almost 100 points. Foreign language spoken at home is the only variable in the regressions that varies substantially depending on whether country fixed effects or country-level variables and interaction terms are used. In regressions including country fixed effects or only country-level variables but no interactions, the coefficient estimate of this variable is considerably lower: the disadvantage of students speaking a foreign language at home shrinks to below 20 points (Table A2 in Annex A). The sensitivity of this result persists in school-level regressions as well (Table A4 in Annex A). Even with this caveat in mind, this result suggests more resources be devoted to reducing language barriers for the affected students.

65. Estimation results also indicate the presence of gender disparities with respect to student outcomes. Girls had lower PISA scores in science and math in comparison with boys, though girls scored higher in reading. Some improvement occurred between 2015 and 2018 as girls improved over time: the negative gap was shrinking for science and math and girls' relative advantage in reading widened considerably. These results are confirmed in some other empirical research (Faria and Portea 2016; Gamazo et al., 2018). To the extent that these gender gaps can be further closed (to the extent that they are the consequence of gender stereotyping rather than inherent biological differences), there is a potential for a further improvement in student performance, with potential implications for aggregate productivity. If girls' average performance in maths and science rose to the average level of boys, and vice versa for reading, then average national PISA scores would be increased by about 3 points with a potential to raise long-run aggregate productivity by around ½ per cent, with larger effects for countries where initial disparities are larger (Figure 10).

Figure 10. Gender disparities in PISA scores

Fractions of a standard deviation across each subject



Note: The bars represent the gender differences in PISA scores in a direction that might be thought to confirm gender stereotypes, namely between boys' less girls' scores for science and math and girls' less boys' scores for reading. These differences are then standardised by dividing by the within-country standard deviation of PISA scores in each subject before the scores are aggregated across subjects.
Source: PISA 2018.

Table 6. Estimation results – socioeconomic background

PISA=f(socioeconomic background, other policies, controls)

	Average	Science	Math	Reading	Minimum	Maximum
Independent variables						
Socioeconomic background (SEC)*	26.086**	25.804**	25.724**	26.729**	-7.7	4.5
Interactions						
SEC X country dummy for above average inequality	6.521**	7.159**	7.451**	4.952**		
School-level standard deviation of SEC	-25.006**	-24.663**	-23.417**	-26.937**	0.0	2.8
Gini coefficient of income inequality**	-2.036**	-1.881**	-3.087**	-1.139**	25.4	51.0
GDP per capita	0.391**	0.756**	-0.343**	0.760**	11.8	107.4
Foreign language spoken at home	-94.244**	-103.236**	-68.963**	-110.533**	0	1
Female student (dummy)	-0.035	-5.405**	-13.735**	19.034**	0	1

Note: * The socioeconomic background variable is derived from three variables related to family background, namely highest parental education , highest parental occupation, and home possessions including books in the home and is demeaned by country.
** Higher value = more inequality. Otherwise as for Table 2.

8. Summarising potential aggregate productivity gains

66. This section summarises the potential gains to aggregate productivity from the various scenarios considered in earlier sections of the paper (Table 7). In most cases these scenarios are not particularly ambitious, considering that only lagging countries catch-up to the median performing country on a particular policy (the only exceptions are accountability and teacher quality where the benchmark is set at

the upper quartile). It should, however, be emphasised again that the full long-run productivity effects will only be realised after several decades once the effect of the change is fully reflected in the human capital of the entire labour force.

67. The largest potential gains in productivity only apply to policy changes relevant to relatively few OECD countries. More specifically, reducing income inequalities in the most unequal OECD countries (Colombia, Costa Rica and Mexico) might induce long-term productivity gains of between 4% to 6%, and raising education spending in low-spending countries (Colombia, Mexico and Türkiye) would boost productivity by between 3% and 4½ per cent. [18]

68. More widespread gains could be achieved from raising ECEC attendance of three-to-five-year-olds, with increases in long-run productivity of between1% and 2% for eight OECD countries and nearly 3% for Türkiye. Similarly, more than half of OECD countries would stand to gain 1% to 2% in the long-run productivity by improving the qualifications of teachers, although this result needs to be recognised more broadly as reflecting the importance of the quality of teachers. On the other hand, reducing the extent of grade repetition and extensive ability grouping implies only modest productivity improvements for only a handful of OECD countries.

69. Adding up the potential productivity gains for individual OECD countries provides useful information on the overall scope for policy reforms that raise educational performance to raise aggregate productivity in individual OECD countries (Table 7). In some countries (such as Chile, Colombia, Costa Rica, Mexico, and Türkiye) that are lagging behind in a number of areas, a package of reforms is estimated to enhance country-level productivity by about 10% or more (final column of Table 7). It should be noted that caution is warranted for adding spending-related productivity improvement to other policy reforms, as spending effects might finance some of these reforms. Other countries including Australia, Greece, Ireland, Lithuania, Portugal, Switzerland or the United States perform well in some areas but are laggards in others, so that implementing policy reforms in areas where they underperform would boost productivity by around 3%. Finally, in a few countries that are close to best practice in most policies (Canada, Norway and Sweden), potential productivity gains from policy reforms are negligible (although this result needs to be qualified where recent data is not available to evaluate the current policy stance).

[18] By comparison, a one-year increase in the quantity of education for an average OECD country with mean years of schooling of 12.5 years would go hand in hand with an increase in productivity by 1.6% (calculations based on Égert et al, 2022).

Table 7. Long-run productivity gains from policy reforms to raise educational performance.

(in per cent)

	Increase early childhood education and care participation	**Improve teacher quality**	**Reduce ability grouping**	**Reduce grade repetition**	**Increase accountability**	**Increase aggregate spending**	**Reduce income inequality**	**Total potential productivity gains***
	Participation rate 3-5y	Share of teachers with master's degree	Share of schools separating students in all subjects	Once in lower secondary	Share of schools posting achievement data publicly	Expenditure per student, below sample median	Gini coefficient	
Scenario converging to:	Median	Upper quartile	Median	Median	Upper quartile	Median	Median	
AUS	0.3	1.4	-	-	0.1	-	0.5	2.3
AUT	0.1	0.6	-	0.1	0.3	-	-	1.2
BEL	-	0.9	0.1	0.4	0.3	-	-	1.7
CAN	-	n.a	0.0	-	0.1	-	0.0	0.1
CHE	1.9	0.4	0.2	0.2	0.3	n.a	0.2	3.2
CHL	0.9	1.8	-	0.1	0.2	2.1	3.8	8.9
COL	n.a	1.8	0.1	1.0	0.2	4.5	6.3	13.9
CRI	n.a	1.5	0.6	0.8	0.3	n.a	5.2	8.4
CZE	0.4	-	-	-	0.2	-	-	0.6
DEU	-	0.0	0.0	0.4	0.3	-	-	0.8
DNK	-	1.7	0.0	-	0.2	-	-	1.8
ESP	-	0.9	0.0	0.9	0.3	-	0.7	2.8
EST	0.0	0.1	-	-	0.2	-	-	0.3
FIN	0.2	-	-	-	0.3	-	-	0.5
FRA	-	1.0	0.0	0.2	0.1	-	-	1.3
GBR	-	1.4	-	-	0.0	-	0.0	1.4
GRC	1.7	1.5	-	-	0.2	-	0.3	3.8
HUN	-	0.3	-	0.0	0.2	1.3	-	1.8
IRL	-	1.2	-	-	0.2	-	-	1.3
ISL	-	1.4	-	-	0.2	-	-	1.7
ISR	-	1.0	0.0	0.1	-	-	1.8	2.9
ITA	0.1	0.3	0.1	0.1	0.2	-	0.8	1.6
JPN	-	n.a	0.0	n.a	0.3	-	0.1	0.4
KOR	-	1.0	-	0.0	0.2	-	-	1.2
LTU	n.a	0.8	0.0	-	0.2	0.8	1.0	2.8
LUX	0.2	0.1	0.2	0.8	0.3	-	0.3	1.7
LVA	n.a	0.7	-	-	0.2	0.3	0.9	2.0
MEX	1.0	1.6	0.1	0.0	0.2	4.0	4.1	11.0
NLD	-	1.2	0.4	0.2	-	-	-	1.7
NOR	-	n.a	0.0	n.a	0.0	-	-	0.0
NZL	0.1	n.a	0.0	-	-	-	n.a	0.1
POL	0.2	-	-	-	0.1	-	-	0.3
PRT	0.1	1.6	-	0.4	0.1	-	0.6	2.8
SVK	0.7	-	-	-	0.1	0.0	-	0.8
SVN	0.0	1.6	-	-	0.1	-	-	1.7
SWE	-	n.a	-	-	0.1	-	-	0.1
TUR	n.a	1.5	0.1	-	-	2.8	2.7	7.2
USA	1.7	0.7	-	0.0	-	-	1.9	4.3
*Average***	*0.3*	*0.9*	*0.1*	*0.2*	*0.2*	*0.4*	*0.8*	*2.9*
*Maximum****	*1.9*	*1.8*	*0.6*	*1.0*	*0.3*	*4.5*	*6.3*	*13.9*

Note: The scenarios are shown in “Scenario converging to” indicate that individual countries are considered to converge to the OECD median, More details and discussion about the scenarios and variables being shocked can be found in sections 3 to 7 in the main text. "-" means that the country is above the convergence criteria i.e., is already a good performer. "n.a." means there is no data for the country in the *PISA* survey or *Education at a Glance* for early childhood education.* Total is the sum of all scenarios for each country; ** Average of all countries including those with 0 effect; *** Maximum represents the maximum effect across all countries.

9. Concluding remarks

70. Using data from the OECD's 2018 *Programme for International Student Assessment* (PISA), this paper attempts to quantify the link between policies and student test scores. In combination with earlier work done in Égert et al. (2022) that established a link between macroeconomic productivity and a measure of human capital combining student test scores and mean years of schooling, it is possible to quantify the effect of policies impacting educational performance on macroeconomic productivity.

71. Results highlight the importance of early childhood education and care (ECEC). Students having previously experienced up to three years of ECEC achieve significantly better PISA scores than those not having attended ECEC. Scaling up this effect and linking it to an improvement in human capital, suggests that a 10-percentage point improvement in the coverage of three-to-five-year-olds in ECEC will eventually raise aggregate productivity by 0.5% in the long run, although it will take several decades before the full macroeconomic effect is achieved.

72. A positive association exists between education spending per student and student test scores though only for spending levels below the OECD median. There may be scope for currently low-spending countries to raise student performance -- especially Mexico and Colombia -- with potential long-term gains equivalent to 4% to 5% for productivity.

73. Estimations identify a positive relation between the share of teachers with a master's degree, one proxy for teacher quality, and PISA scores with potential long-run productivity gains reaching 1% to 2% for lagging countries. As teacher qualification is strongly correlated with other measures of teacher quality, this result should be seen as a policy, among others, to improve teacher quality.

74. PISA results are worse in schools that separate students by ability and achievement in all subjects (with much larger effects for students from a lower socioeconomic background or those speaking a foreign language at home) and for students repeating a grade. At the country level, however, possible productivity gains tend to be modest in most OECD countries due to the low incidence of ability grouping and grade repetition.

75. Greater accountability, both in terms of schools posting achievement data publicly and especially the presence of externally administered exit exams in secondary education, correlate with higher PISA scores. Some schools could raise long-run productivity by as much as 0.3% by posting achievement data publicly. Also, confirming previous results in the literature, greater autonomy for schools only has a modest positive effect on student performance even when combined with accountability.

76. Reducing income inequality can raise educational performance and so raise aggregate productivity. Reducing the highest levels of inequality - especially in Colombia, Mexico and Costa Rica - to the OECD median could raise long-run aggregate productivity by 4% to 6%. Furthermore, the strong effect of students' socioeconomic background on student performance is magnified in countries with higher inequality.

77. Results suggest that girls perform better in reading, whereas boys have an edge in maths and science. Levelling up of scores -- so that girls catch-up with boys in maths and science and boys catch-up with girls in reading – is estimated to raise productivity by around ½ per cent for an OECD country with an average level of gender disparity in learning outcomes.

REFERENCES

Akabayashi H, Tanaka R (2013), “Long-term effects of preschooling on educational attainments”, *Keio/Kyoto Joint Global COE Discussion Paper Series*.

Akiba, M., Chiu, Y., Shimizu, K. and Liang, G. (2012). “Teacher Salary and National Achievement: A Cross-National Analysis of 30 Countries”, *International Journal of Educational Research*, 53,171-181.

Attanasio, O.P. (2015), “The determinants of human capital formation during the early years of life: theory, measurement and policies“, *Journal of the European Economic Association*, voil. 13 (6),pp. 949–997.

Avvisati, F. and P. Givord (2023), "The learning gain over one school year among 15-year-olds: An international comparison based on PISA", *Labour Economics*.

Baker, M., J. Gruber, K. Milligan, (2008), “Universal Child Care, Maternal Labor Supply, and Family Well-Being”, *Journal of Political Economy*, *116*(4), 709–745.

Balladares, J. and M. Kankaraš (2020), "Attendance in early childhood education and care programmes and academic proficiencies at age 15", *OECD Education Working Papers*, No. 214, OECD Publishing, Paris.

Barra, C. and M. Boccia (2022), “What matters in educational performance? Evidence from OECD and non-OECD countries, *Quality and Quantity*, 56, 4335-4394

Berg A. J., and J. D. Ostry, C. G. Tsangarides and Y. Yakhshilikov (2018), “Redistribution, Inequality and Growth: New Evidence”, *Journal of Economic Growth* 23, 259-305, September.

Berger L.M., L. Panico and A. Solaz, “The Impact of Center-Based Childcare Attendance on Early Child Development: Evidence From the French Elfe Cohort”, *Demography* 58 (2), 419–450.

Bigras, N., & Lemay, L. (2012). Petite enfance, services de garde éducatifs et développement des enfants : État des connaissances. Presses Universitaires du Québec.

Bingley, P., V.M. Jensen and I Walker (2005), “The effect of school class size on length of post-compulsory education: some cost-benefit analysis“, *IZA Discussion Paper* No. 1605.

Boarini, R. (2009), “Towards better schools and more equal opportunities for learning in Italy”, *OECD Economics Department Working Papers*, No. 727, OECD Publishing, Paris.

Boarini, R., and Lüdemann, E., (2009), The role of teacher compensation and selected accountability policies for learning outcomes: An empirical analysis for OECD countries. *OECD Journal: Economic Studies*, 211-230.

Bonacini, L., I. Brunetti and G. Giovanni, (2021), “Choose the school, choose the performance. New evidence on the determinants of student performance in eight European countries*”, RECent Centrer for Economic Research Working Paper* No. 150.

Braga, M., D., Checchi and E. Meschi, (2013), “Educational policies in a long-run perspective“, *Economic Policy*, 28 (73), 45–100.

Browning, M. and E. Heinesen (2007), “Class size, teacher hours and educational attainment“, Scandinavian Journal of Economics, 109(2), 415–438.

Cascio, E. U. (2009). Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools. *The Journal of Human Resources*, *44*(1), 140–170.

Cerra, V., R. Lama and N. Loayza (2021), “Links between growth, inequality and poverty: A survey, *IMF Working Paper* No.WP/21/68.

Chingos, M. M. (2013), “Class size and student outcomes: Research and policy implications”, *Journal of Policy Analysis and Management*, 32(2), 411-438.

Cingano, F. (2014), “Trends in Income Inequality and its Impact on Economic Growth”, *OECD Social, Employment and Migration Working Papers* No. 163, OECD Publishing, Paris.

Cunha, F., J.J. Heckman, L. J. Lochner, and D. V. Masterov (2006). "Interpreting the Evidence on Life Cycle Skill Formation", In *Handbook of the Economics of Education*, ed. Eric A. Hanushek and Frank Welch, 697-812. Amsterdam: North-Holland: Elsevier.

De Philippis, M. and F. Rossi (2021), "Parents, Schools and Human Capital Differences Across Countries", Journal of the European Economic Association.

Denny, K. and V. Oppedisano, (2013), “The surprising effect of larger class sizes“, Labour Economics, 23, 57–65.Dolton, Peter and O. 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.

Dong, F. and M. Cornachione Kula (2023) Digital device use and scientific literacy: an examination using Programme for International Student assessment (PISA) 2015 data, Education Economics, 31:3, 288-312.

Égert, B. and P. Gal (2016), “The quantification of structural reforms in OECD countries: A new framework”, OECD Economics Department Working Paper No. 1354.

Égert, B., C.de la Maisonneuve and D. Turner (2022), “A new macroeconomic measure of human capital exploiting PISA and PIAAC: Linking education policies to productivity”*, OECD Economics Department Working Papers*, No. 1709, OECD Publishing, Paris.

Égert, B., J. Botev and D. Turner (2020), "The contribution of human capital and its policies to per capita income in Europe and the OECD," *European Economic Review, 129©.*

Faria, S. and M.C. Portela (2016), “Student performance in mathematics using PISA-2009 data for Portugal”, *Católica Porto Business School, Universidade Católica Portuguesa, Working Paper* No. 1. https://EconPapers.repec.org/RePEc:cap:mpaper:012016

Felfe, C., Nollenberger, N., and Rodríguez-Planas, N. (2015). Can’t Buy Mommy’s Love? Universal Childcare and Children’s Long-Term Cognitive Development. *Journal of Population Economics*, 28, 393–422

Fredriksson, P., B. Oeckert and H. Oosterbeek (2013), “Long-term effects of class size”, *Quarterly Journal of Economics*, 128(1), 249-285.

Fuchs, T. and L. Woessmann (2007), “What Accounts for International Differences in Student Performance? A Re-examination Using PISA Data”, *Empirical Economics*, 32 (2007), 433-464.

Fullard, J. (2021), "Relative wages and pupil performance, evidence from TIMSS," *ISER Working Paper Series 2021-07*, Institute for Social and Economic Research.

Galor, O. and O. Moav (2004), “From Physical to Human Capital Accumulation: Inequality and the Porcess of Development”, *The Review of Economic Studies* 71(4), pp. 101-1026.

Gamazo, A., F. Martínez-Abad, S. Olmos-Migueláñez and M.J. Rodríguez-Conde, (2018), "Assessment of factors related to school effectiveness in PISA 2015". *A multilevel analysis. Revisit. Educ.*, 379, 56–84.

Glover, D. and R. Levacic (2020), "Funding education – public and private systems", in *Educational Resource management: An international perspective*, UCL Press, 16-36.

Gomendio, M. (2023), PISA: Mission failure, *Standards, Testing and Accountability*, 23(2), *PISA: Mission Failure - Education Next.*

Goos, M., B.M. Schreier, H.M.E. Knipprath, B. De Fraine, J. Van Damme and U. Trautwein (2013), "How Can Cross-Country Differences in the Practice of Grade Retention Be Explained? A Closer Look at National Educational Policy Factors", *Comparative Education Review*, 57(1), 54-84.

Goos, Mieke; Schreier, Brigitte Maria; Knipprath, Heidi Maria Eduard; De Fraine, Bieke; Van Damme, Jan; Trautwein, Ulrich

Goux, D. and E. Maurin (2010), "Public school availability for two-year olds and mothers' labour supply," *Labour Economics*, 17(6), 951-962.

Guillemette, Y., A. Kopoin, D. Turner and A. De Mauro (2017), "A revised approach to productivity convergence in long-term scenarios", *OECD Economics Department Working Papers* No. 1385.

Hanushek E. A., Piopiunik, M. and S. Wiederhold (2019), "The Value of Smarter Teachers: International Evidence on Teacher Cognitive Skills and Student Performance", *Journal of Human Resources*, 54(4), 857-899.

Hanushek, E. A. and L. Woessmann (2006), "Does Early Tracking Affect Educational Inequality and Performance? Differences-in[1]Differences Evidence across Countries", *Economic Journal*, 116 (510), 63–76.

Hanushek, E. A. and L. Woessmann (2011), "The Economics of International Differences in Educational Achievement", *Handbook of the Economics of Education*, 3, 89-200.

Hanushek, E. and J. A. Luque (2003). "Efficiency and Equity in Schools Around the World" *Economics of Education Review,* 22, pp.481-502.

Heckman, J. J., R. Pinto and P. Savelyev (2013), "Understanding the Mechanisms, through which an Influential Early Childhood Program Boosted Adult Outcomes", *American Economic Review*, Vol. 103 (6), pp. 2052-66.

Heckman, J.J. (2008), "Schools, skills and synapses", *NBER Working Papers* No. 14064.

Hippe, R., M. Jakubwski and L. Araujo (2018), Regional inequalities in PISA: the case of Italy and Spain, *JRC Technical Reports*, EUR 28868; doi:10.2760/495702.

Hodgen, E. (2007), Early childhood education and young adult competencies at age 16", *Technical report 2 from the age 16 phase of the longitudinal Competent Children*, Competent Learners study. Wellington: Ministry of Education.

Hong, J. (2015), "Effects of Education Policies and Institutions on Student performance", *Seoul Journal of Economics*, 28(1), 85-106.

Ikeda, M. and E. García (2014), "Grade repetition: A comparative study of academic and non-academic consequences", *OECD Journal: Economic Studies*, 2013/1, 269-315.

Jacob, B. and L. Lefgren (2009), "The effect of grade retention on high school completion", *American Economic Journal: Applied Economics*, 1(3) 33-58.Jaumotte, F. (2003), "Female labour force participation: past trends and main déterminants in OECD countries", *OECD Economics Department Working Papers* No. 376.

Jepsen, C. and S. Rivkin (2009), "Class size reduction and student achievement: The potential tradeoff between teacher quality and class size", *Journal of Human Resources*, 44(1), 223-250.

Jones, C. (2016), "The Facts of Economic Growth", in Handbook of Macroeconomics, https://doi.org/10.1016/bs.hesmac.2016.03.002.

Juerger, H. and K .Schneider (2004), "International Differences in Student Achievement: An Economic Perspective", *German Economic Review,* 5(3), 357-380.

Lavrijsen, J. and I. Nicaise (2015), "New empirical evidence on the effect of educational tracking on social inequalities in reading achievement", *European Educational Research Journal*, 14(3-4), 206-221.

Leuven, E. and H. Ooesterbeek (2018), "Class size and student outcomes in Europe", Analytical Report No. 33 Prepared for the European Commission, European Expert Economics of Education.

Lian, Q., C. Yu, X. Tu, M. Deng, T. Wang, Q. Su and X. Zou (2021), "Grade repetition and bullying victimmisation in adolescents: A global cross-sectional study of the Program for International Student Assessment (PISA) data from 2018", *PLOS Medicine*, November, 1-17.

Murat, M. and L. Bonacini (2020), "Coronavirus pandemic, remote learning and education inequalities". GLO Discussion Paper No. 679.

OECD (2005), *Teachers matter: attracting, developing and retaining effective teachers*, OECD Publishing, Paris.

OECD (2018a), *PISA 2018 results (Volume III): What school life means for students' lives*, OECD Publishing, Paris.

OECD (2018b), "*Responsive School Systems*", OECD Publishing, Paris.

OECD (2018c), *PISA 2018 results: Effective policies, successful schools, Volume V*, OECD Publishing, Paris.

OECD (2018d), *PISA 2018 results (Volume II): Where all students can succed*, OECD Publishing.

OECD (2021), Starting Strong VI: Supporting Meaningful Interactions in Early Childhood Education and Care, Starting Strong, OECD Publishing.

OECD (2022a), *The importance of human capital for economic outcomes, Chapter 1 in: Value for Money in School Education: Smart Investments Outcomes, Equal Opportunities*, OECD Publishing, Paris.

OECD (2022b), "Governing and distributing school funding: Effectively connecting resources and learning", Chapter 3 in: *Value for Money in School Education: Smart Investments Outcomes, Equal Opportunities*, OECD Publishing, Paris.

OECD (2022c), Using school funding to achieve both efficiency and equity in education, Chapter 4 in: *Value for Money in School Education: Smart Investments Outcomes, Equal Opportunities*, OECD Publishing, Paris.

OECD (2023a), "Proposal for an Action Plan to Reduce Early Leaving in Spain", *OECD Education Policy Perspectives*, OECD Publishing, Paris.

OECD (2023b), *OECD Economic Surveys: Spain 2023*, OECD Publishing, Paris, forthcoming.

OECD (2023c), Education at a Glance 2023: OECD Indicators, OECD Publishing, Paris, https://doi.org/10.1787/e13bef63-en.

Phillips, D. A. et al. (2017), Puzzling It Out: The Current State of Scientific Knowledge on Pre-Kindergarten Effects A Consensus Statement, Brookings, consensus-statement_final.pdf (brookings.edu)

Quintano, C., R. Castellano and S. Longobardi (2009), “L'influenza dei fattori socio economici sulle competenze degli studenti italiani. Un'analisi multilevel dei dati PISA 2006“, *Rivista di Economia e Statistica del Territorio*, 2, 1-43.

Schneeweis, N., and R. Winter-Ebmer (2007), “Peer effects in Austrian schools”, *Empirical Economics*, 32(2), 387–409.

Schuetz, G., (2009), “Does the Quality of Pre-primary Education Pay Off in Secondary School? An International Comparison Using PISA 2003”, *ifo Working Paper*, No. 68.

Slot, P. (2018). Structural characteristics and process quality in early childhood education and care : A literature review. OECD Education Working Paper.

Smidova, Z. (2019), “Educational Outcomes: A Literature Review of Policy Drivers from a Macroeconomic Perspective”, *OECD Economics Department Working Papers*, No. 1577.

Strello, A., R. Strietholt, I. Steinmann and C. Siepmann (2021), “Early Tracking and Different Types of Inequalities in Achievement: Difference-in-Differences Evidence from 20 Years of Large-scale Assessments”, *Educational Assessment, Evaluation and Accountability*, 33, 139-167.

Terrin, E. and M. Triventi (2022), “The effect of school tracking on student achievement and inequality: A meta-analysis”, *Review of Educational Research*, 93(2), 236-274.

Thévenon, O. (2013), “Drivers of female labour force participation in the OECD”, *OECD Social, Employment and Migration Working Papers* No. 145.

Torres, R. (2021), "Does test-based school accountability have an impact on student achievement and equity in education?, A panel approach using PISA", *OECD Education Working Papers*, No. 250.

VerBruggen, R. (2019), “Challenging the ‘Heckman Curve’: Are public investments in kids really the best?”, Institute for Family Research, Challenging the ‘Heckman Curve’: Are Public Investments in Kids Really the Best? | Institute for Family Studies (ifstudies.org)

Weiland, C. et al. (2023), "Lottery-based evaluations of early education programs: Opportunities and challenges for building the next generation of evidence", *NBER Working Papers*, No. 30970.

West, M.R. and L. Woessmann (2010), “Every Catholic Child in a Catholic School: Historical Resistance to State Schooling, Contemporary School Competition, and Student Achievement across Countries”, *Economic Journal*, 120(546), 229-255.

Whitehurst, G. I. R. (2017). “Rigorous preschool research illuminates policy (and why the Heckman equation may not compute)”, Brookings Institution, Rigorous preschool research illuminates policy (and why the Heckman Equation may not compute) (brookings.edu).

Woessmann, L. (2016), “The Importance of School Systems: Evidence from International Differences in Student Achievement”, *Journal of Economic Perspectives*, 30(1), 3-32.

Woessmann, L. (2018), "Central exit exams improve student outcomes", *IZA World of Labor, 419*.

Annex A. Full regression results

This annex presents the full regression results and a comparison of the results between 2015 and 2018 as well as some sensitivity analysis to the baseline regressions. Table A1 shows the full results and presents a comparison of the results in 2015 and 2018 for the three subjects (science, math and reading) and their average. Table A2 presents alternative specifications for 2015 and 2018 for the average PISA score. The first specification includes individual and school level variables only with fixed effects but no interactions. In the second one, fixed effects are replaced by country-level variables. In the third one, interaction terms are added. In table A3, the country coverage is maximised in 2015 and 2018 by dropping selected variables. In a second step, the very same specification is estimated for the two years. Finally, table A4 presents the results at the school level for 2018.

Table A1. Full regression results reported in Section 3 to 7

Dependent variable: PISA score	2015				2018			
	Average (1)	Science (2)	Math (3)	Reading (4)	Average (5)	Science (6)	Math (7)	Reading (8)
Part A. Early childhood education and care (ECEC)								
1 year in ECE	28.021**	26.400**	26.812**	30.851**	33.991**	36.649**	31.518**	33.807**
2 to 3 years in ECE	34.117**	31.804**	34.395**	36.151**	34.035**	33.821**	34.363**	33.923**
4 to 5 years in ECE	30.035**	28.500**	28.943**	32.662**	25.464**	25.836**	25.203**	25.353**
Part B. Resources devoted to education								
School size	0.006**	0.006**	0.005**	0.006**	0.009**	0.008**	0.009**	0.010**
Average class size	0.162**	0.034	0.102	0.350**	0.408**	0.376**	0.417**	0.432**
Staff shortage	-3.417**	-4.079**	-2.897**	-3.275**	-3.119**	-2.844**	-3.349**	-3.162**
Public school	-15.092**	-13.342**	-13.532**	-18.401**	-8.197**	-8.356**	-7.151**	-9.083**
Teachers quality								
Teachers with a Master degree	23.791**	22.962**	23.772**	24.640**	15.514**	13.770**	17.352**	15.422**
Education spending below sample median	0.019**	0.020**	0.017**	0.020**	0.008**	0.006**	0.008**	0.008**
Education spending above sample median	-0.002**	-0.002**	-0.001**	-0.003**	-0.000*	-0.001**	0.001**	-0.001**
Part C. Indirect resources: Tracking and grade repetition								
Schools with ability grouping	-7.740**	-6.972**	-5.909**	-10.340**	-7.005**	-8.183**	-5.688**	-7.143**
First age of tracking	-0.239	-0.863**	-0.283	0.428	-0.408	-0.132	-1.913**	0.820**
Repeated once in primary	-52.077**	-49.984**	-50.759**	-55.489**	-54.293**	-49.738**	-52.539**	-60.603**
Repeated twice or more in primary	-51.002**	-49.548**	-51.405**	-52.052**	-44.620**	-32.632**	-49.906**	-51.323**
Repeated once in lower secondary	-38.163**	-38.021**	-38.356**	-38.113**	-41.013**	-36.144**	-42.421**	-44.473**
Repeated twice or more in lower secondary	-24.508**	-27.730**	-17.957**	-27.837**	-24.600**	-23.497**	-26.050**	-24.253**
Interactions								
Schools with ability grouping x socioeconomic background below average	-7.250**	-7.020**	-5.844**	-8.885**	-9.298**	-9.505**	-10.221**	-8.168**
Schools with ability grouping x foreign language spoken at home	-3.808	-1.938	-2.987	-6.500	-15.357**	-18.251**	-16.845**	-10.975**
First age of tracking x foreign language spoken at home	2.617**	2.593**	3.091**	2.167**	5.321**	5.624**	4.047**	6.293**
Repeated once in lower secondary x Foreign language spoken at home	-3.192	-1.059	-2.460	-6.056	-6.989*	-10.244**	-0.053	-10.671**
Repeated 2 or + in lower secondary x Foreign language spoken at home	-12.000	-0.157	-17.380*	-18.464*	-52.655**	-41.905**	-51.863**	-64.198**
Part D. School governance : accountability and autonomy								
Accountability								
Posting achievement data publicly	3.233**	4.814**	1.454	3.430**	4.167**	4.785**	3.531**	4.184**
External exit exam	2.587**	2.537**	1.962**	3.261**	4.509**	4.121**	4.560**	4.845**
Autonomy								
Budgeting	1.084**	1.196**	0.840*	1.215**	n.a	n.a	n.a	n.a
Staffing	-2.783**	-2.471**	-3.099**	-2.779**	n.a	n.a	n.a	n.a
Curriculum	2.601**	2.502**	2.558**	2.742**	n.a	n.a	n.a	n.a
Interactions								
Accountability x Staff autonomy	1.026*	1.881**	0.640	0.557	n.a	n.a	n.a	n.a
Part E. Socioeconomic background and income inequality								
Socioeconomic background (country demeaned)	21.983**	23.202**	21.994**	20.752**	26.086**	25.804**	25.724**	26.729**
Dispersion in socioeconomic background within schools	-10.405**	-7.712**	-13.375**	-10.127**	-25.006**	-24.663**	-23.417**	-26.937**
Income inequality - Gini index	-0.320**	0.177	-1.760**	0.624**	-2.036**	-1.881**	-3.087**	-1.139**
GDP per capita	-0.366**	-0.251**	-0.734**	-0.112	0.391**	0.756**	-0.343**	0.760**
Being a girl	-5.050**	-12.070**	-17.021**	13.940**	-0.035	-5.405**	-13.735**	19.034**
Foreign language spoken at home	-57.762**	-61.527**	-60.878**	-50.881**	-94.244**	-103.236**	-68.963**	-110.533**
Interactions								
Socioeconomic background (below average) x Teachers with a Master degree	13.182**	17.116**	13.070**	9.359**	11.343**	11.083**	13.168**	9.776**
Socioeconomic background x low inequality	13.184**	12.077**	13.779**	13.696**	6.521**	7.159**	7.451**	4.952**
Socioeconomic background below average x class size	-0.089*	-0.197**	-0.124**	0.054	-0.352**	-0.399**	-0.376**	-0.282**
Part F. Miscellaneous								
Age	13.442**	13.572**	12.382**	14.372**	14.733**	13.850**	14.694**	15.655**
Teachers picking on pupils	-14.457**	-14.621**	-13.546**	-15.204**	n.a	n.a	n.a	n.a
Constant	179.886**	164.309**	281.417**	93.931**	233.947**	237.793**	322.755**	141.292**
R-squared	0.330	0.312	0.325	0.291	0.317	0.282	0.324	0.278
Number of observations	110742	110742	110742	110742	141529	141529	141529	141529
Number of countries	25	25	25	25	26	26	26	26
Country fixed effects	No	No	No	No	No	No	No	No

Note: * and ** denote statistical significance at the 10% and 5% levels, based on heterogeneity-robust standard errors. Countries included in the estimation sample are: Australia, Belgium, Chile, Colombia, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Korea, Latvia, Luxembourg, Mexico, Poland, Portugal, Slovak Republic, Spain, Türkiye, United Kingdom and United States in 2015 and Australia, Chile, Colombia, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Italy, Korea, Latvia, Luxembourg, Mexico, Netherlands, Poland, Portugal, Slovak Republic, Slovenia, Spain, Türkiye, United Kingdom and United States in 2018.

Table A2. Alternative specifications – Excluding country-level variables and interaction terms

Student-level regressions, 2015 and 2018

Dependent variable: average PISA score (science, math, reading)	2015			2018		
	(1)	(2)	(3)	(4)	(5)	(6)
	Indvidual- and school level variables only	(1) + country-level variables	(2) + interaction terms	Indvidual- and school level variables only	(4) + country-level variables	(5) + interaction terms
Part A. Early childhood education and care (ECEC)						
1 year in ECE	28.836**	28.403**	28.021**	33.780**	34.283**	33.991**
2 to 3 years in ECE	35.207**	34.358**	34.117**	39.824**	34.597**	34.035**
4 to 5 years in ECE	30.425**	30.300**	30.035**	30.540**	27.120**	25.464**
Part B. Resources devoted to education						
School size	0.004**	0.006**	0.006**	0.009**	0.009**	0.009**
Average class size	0.346**	0.191**	0.162**	0.557**	0.590**	0.408**
Staff shortage	-4.060**	-3.385**	-3.417**	-3.273**	-3.569**	-3.119**
Public school	-9.138**	-14.796**	-15.092**	-3.775**	-7.722**	-8.197**
Teachers quality						
Teachers with a Master degree	18.277**	20.101**	23.791**	18.505**	11.732**	15.514**
Education spending below sample median		0.019**	0.019**		0.007**	0.008**
Education spending above sample median		-0.002**	-0.002**		-0.000	-0.000*
Part C. Indirect resources: Tracking and grade repetition						
Schools with ability grouping for all subjects	-5.933**	-5.061**	-7.740**	-1.837	-4.570**	-7.005**
First age of tracking		-0.197	-0.239		0.009	-0.408
Repeated once in primary	-52.443**	-52.807**	-52.077**	-54.261**	-54.838**	-54.293**
Repeated twice or more in primary	-52.351**	-51.930**	-51.002**	-47.992**	-46.189**	-44.620**
Repeated once in lower secondary	-41.062**	-39.034**	-38.163**	-46.077**	-43.615**	-41.013**
Repeated twice or more in lower secondary	-34.413**	-30.307**	-24.508**	-39.083**	-32.772**	-24.600**
Interactions						
Schools with ability grouping x socioeconomic background below average			-7.250**			-9.298**
Schools with ability grouping x foreign language spoken at home			-3.808			-15.357**
First age of tracking x foreign language spoken at home			2.617**			5.321**
Repeated once in lower secondary x Foreign language spoken at home			-3.192			-6.989*
Repeated 2 or + in lower secondary x Foreign language spoken at home			-12.000			-52.655**
Part D. School governance : accountability and autonomy						
Accountability						
Posting achievement data publicly	5.628**	3.018**	3.233**	3.412**	3.743**	4.167**
External exit exam		2.335**	2.587**		4.245**	4.509**
Autonomy						
Budgeting	1.346**	0.851*	1.084**	n.a	n.a	n.a
Staffing	-0.602	-2.311**	-2.783**	n.a	n.a	n.a
Curriculum	-0.169	2.734**	2.601**	n.a	n.a	n.a
Interactions						
Accountability x Staff autonomy			1.026*	n.a	n.a	n.a
Part E. Socioeconomic background and income inequality						
Socioeconomic background (country demeaned)	23.422**	22.701**	21.983**	24.960**	25.191**	26.086**
Dispersion in socioeconomic background within schools	-5.086	-10.367**	-10.405**	-31.028**	-24.380**	-25.006**
Income inequality - Gini index		-0.277**	-0.320**		-1.966**	-2.036**
GDP per capita		-0.380**	-0.366**		0.386**	0.391**
Being a girl	-5.335**	-5.038**	-5.050**	0.099	0.267	-0.035
Foreign language spoken at home	-18.011**	-18.692**	-57.762**	-16.596**	-18.675**	-94.244**
Interactions						
Socioeconomic background (below average) x Teachers with a Master degree			13.182**			11.343**
Socioeconomic background x low inequality			13.184**			6.521**
Socioeconomic background below average x class size			-0.089*			-0.352**
Part F. Miscellaneous						
Age	12.849**	13.344**	13.442**	15.705**	14.606**	14.733**
Teachers picking on pupils	-14.887**	-14.599**	-14.457**	n.a	n.a	n.a
Constant	296.534**	180.169**	179.886**	211.358**	224.066**	233.947**
R-squared	0.337	0.330	0.330	0.320	0.308	0.317
Number of observations	114355	114355	110742	144215	144215	141529
Number of countries	25	25	25	26	26	26
Country fixed effects	Yes	No	No	Yes	No	No

Note: * and ** denote statistical significance at the 10% and 5% levels, based on heterogeneity robust standard errors.

Table A3. Alternative regressions – Sample maximisation and equalisation

	Sample maximisation		Sample equalisation	
Dependent variable: Average PISA score	2015	2018	2015	2018
	(1)	(2)	(3)	(4)
Part A. Early childhood education and care (ECEC)				
1 year in ECEC	28.741**	33.396**	31.398**	34.491**
2 to 3 years in CECE	34.913**	38.979**	40.341**	35.859**
4 to 5 years in ECEC	30.150**	28.778**	36.413**	26.912**
Part B. Resources devoted to education				
School size	0.004**	0.010**	0.005**	0.009**
Average class size	0.322**	0.387**	0.360**	0.240**
Staff shortage	-3.987**	-2.917**	-3.690**	-3.475**
Public school	-9.674**	-4.486**	-11.855**	-5.065**
Teachers quality				
Teachers with a Master degree	21.567**	22.515**	23.794**	23.515**
Education spending below sample median			0.019**	0.014**
Education spending above sample median			-0.002**	0.001**
Part C. Indirect resources: Tracking and grade repetition				
Schools with ability grouping for all subjects	-8.783**	-3.929**	-8.891**	-6.353**
First age of tracking			0.740*	-3.781**
Repeated once in primary	-51.713**	-53.450**	-55.406**	-58.319**
Repeated twice or more in primary	-51.697**	-46.505**	-45.839**	-41.942**
Repeated once in lower secondary	-39.499**	-43.112**	-43.394**	-45.919**
Repeated twice or more in lower secondary	-28.682**	-32.132**	-35.801**	-20.880**
Interactions				
Schools with ability grouping x socioeconomic background below average	-7.846**	-8.870**	-7.935**	-10.594**
Schools with ability grouping x foreign language spoken at home	-5.665	-18.697**	-3.788	-19.677**
First age of tracking x foreign language spoken at home			3.247**	4.825**
Repeated once in lower secondary x Foreign language spoken at home	-4.496	-7.932*	-2.522	-3.963
Repeated 2 or + in lower secondary x Foreign language spoken at home	-10.342	-42.472**	-36.082**	-68.037**
Part D. School governance : accountability and autonomy				
Accountability				
Posting achievement data publicly	5.623**	3.985**	1.821*	2.644**
External exit exam			0.918**	5.997**
Autonomy				
Budgeting	1.499**	n.a		
Staffing	-0.892	n.a		
Curriculum	-0.155	n.a		
Interactions				
Accountability x Staff autonomy	0.803	n.a		
Part E. Socioeconomic background and income inequality				
Socioeconomic background (country demeaned)	22.786**	25.722**	22.109**	26.336**
Dispersion in socioeconomic background within schools	-5.253*	-32.732**	-8.001**	-28.214**
Income inequality - Gini index			-1.178**	-0.572**
GDP per capita			-0.351**	-0.656**
Being a girl	-5.350**	-0.035	-0.430	0.546
Foreign language spoken at home	-17.114**	-13.383**	-68.627**	-85.408**
Interactions				
Socioeconomic background (below average) x Teachers with a Master degree	14.018**	11.767**	13.220**	8.314**
Socioeconomic background x low inequality	11.794**	7.237**	13.625**	7.066**
Socioeconomic background below average x class size	-0.069	-0.357**	-0.125**	-0.442**
Part F. Miscellaneous				
Age	13.038**	15.848**	14.243**	13.261**
Teachers picking on pupils	-14.699**	n.a		
Constant	294.845**	214.167**	151.242**	255.175**
R-squared	0.338	0.327	0.302	0.316
Number of observations	121799	154179	101317	121053
Number of countries	28	29	22	22
Country fixed effects	Yes	Yes	No	No

Note: * and ** denote statistical significance at the 10% and 5% levels, based on heterogeneity robust standard errors. Sample maximisation: dropping the variables school autonomy and teachers picking on student to increase country coverage (additional three countries added for 2018 include: Costa Rica, Lithuania, and Switzerland). Sample equalisation: estimating the same specification for 2015 and 2018.

Table A4. Regression results at the school level, 2018

Dependent variable: PISA score	Average			Science			Math			Read		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Part A. Early childhood education and care (ECEC)												
1 year in ECE	59.23**	65.82**	57.49**	59.08**	67.98**	59.63**	63.14**	65.40**	56.88**	55.46**	64.08**	55.95**
2 to 3 years in ECE	53.55**	25.49**	20.88**	51.11**	18.79**	14.25**	57.92**	38.77**	33.73**	51.63**	18.89**	14.66**
4 to 5 years in ECE	31.66**	22.15**	12.82**	32.12**	19.72**	10.26**	34.10**	26.51**	17.16**	28.75**	20.21**	11.03**
Part B. Resources devoted to education												
School size	0.30**	0.31**	0.35**	0.25**	0.24**	0.29**	0.27**	0.30**	0.33**	0.38**	0.38**	0.42**
Average class size	0.18**	0.33**	0.26**	0.17**	0.36**	0.26**	0.20**	0.32**	0.27**	0.18**	0.31**	0.25**
Staff shortage	-0.46	-0.92**	-0.82**	-0.26	-0.78*	-0.62	-0.82*	-1.19**	-1.13**	-0.30	-0.80*	-0.69
Public school	14.49**	8.60**	7.22**	14.46**	7.80**	6.41**	15.07**	9.54**	8.02**	13.93**	8.46**	7.21**
Teachers quality												
Teachers with a Master degree	11.12**	6.97**	13.46**	11.33**	5.72**	11.71**	11.67**	8.30**	15.22**	10.35**	6.89**	13.45**
Education spending below sample median		0.01**	0.01**		0.01**	0.01**		0.01**	0.01**		0.01**	0.01**
Education spending above sample median		-0.00**	-0.00**		-0.00**	-0.00**		0.00**	0.00**		-0.00**	-0.00**
Part C. Indirect resources: Tracking and grade repetition												
Schools with ability grouping for all subjects	-1.71	-3.40**	-2.98*	-2.30*	-4.30**	-4.56**	-2.30	-2.67**	-1.27	-0.52	-3.24**	-3.12*
First age of tracking		0.37	-0.37		0.55*	-0.29		-0.90**	-1.48**		1.47**	0.67**
Repeated once in primary	-58.05**	-63.59**	-63.57**	-56.97**	-61.76**	-61.30**	-61.43**	-66.27**	-66.56**	-55.74**	-62.75**	-62.85**
Repeated twice or more in primary	-55.24**	-23.37	-58.03**	-27.56	3.15	-31.90	-58.80**	-32.03	-68.19**	-79.35**	-41.23**	-74.00**
Repeated once in lower secondary	-81.16**	-65.75**	-60.65**	-74.93**	-53.72**	-50.07**	-84.97**	-67.32**	-61.82**	-83.58**	-76.21**	-70.05**
Repeated twice or more in lower secondary	-38.50**	-18.13*	-5.26	-44.96**	-21.59*	-10.73	-23.62**	-4.01	6.42	-46.92**	-28.78**	-11.48
Interactions												
Schools with ability grouping x socioeconomic background below average			-9.79**			-11.63**			-12.19**			-5.53
Schools with ability grouping x foreign language spoken at home			-30.81**			-27.92**			-47.89**			-16.60*
First age of tracking x foreign language spoken at home			9.19**			10.38**			6.88**			10.31**
Repeated once in lower secondary x Foreign language spoken at home			-3.99			7.11			-5.97			-13.10
Repeated 2 or + in lower secondary x Foreign language spoken at home			-56.67			-40.08			-47.47*			-82.44*
Part D. School governance : accountability and autonomy												
Accountability												
Posting achievement data publicly	0.95	1.07	0.91	1.51*	1.91**	1.70*	0.61	0.57	0.50	0.73	0.72	0.53
External exit exam		1.36**	1.42**		0.64**	0.74**		2.06**	2.09**		1.38**	1.44**
Part E. Socioeconomic background and income inequality												
Socioeconomic background (country demeaned)	54.99**	55.28**	54.62**	53.37**	54.05**	53.34**	53.90**	53.76**	53.27**	57.71**	58.04**	57.23**
Dispersion in socioeconomic background within schools	-8.46**	-5.46**	-0.89	-10.32**	-6.59**	-1.91	-7.18**	-3.86	0.82	-7.87**	-5.93**	-1.58
Income inequality - Gini index		-2.33**	-2.48**		-2.20**	-2.34**		-3.43**	-3.58**		-1.36**	-1.52**
GDP per capita		0.59**	0.54**		0.94**	0.89**		-0.02	-0.09		0.85**	0.82**
Being a girl	15.81**	17.26**	14.33**	8.44**	10.27**	7.27**	-1.06	0.36	-2.58	40.05**	41.15**	38.28**
Foreign language spoken at home	-7.94**	-20.09**	-150.80**	-15.30**	-28.40**	-178.27**	0.58	-10.14**	-103.98**	-9.09**	-21.74**	-170.14**
Interactions												
Socioeconomic background (below average) x Teachers with a Master degree			39.24**			36.37**			41.75**			39.59**
Socioeconomic background x low inequality			29.14**			30.82**			29.66**			26.93**
Socioeconomic background below average x class size			-0.07			-0.16			-0.01			-0.04
Part F. Miscellaneous												
Age	24.96**	19.01**	19.29**	25.72**	22.44**	22.62**	26.94**	18.21**	18.69**	22.22**	16.37**	16.57**
Constant	47.13	151.81**	166.04**	47.79	93.51	111.70	20.03	246.35**	254.60**	73.57	115.57	131.81*
R-squared	0.731	0.704	0.724	0.711	0.675	0.697	0.737	0.714	0.734	0.718	0.692	0.709
Number of observations	144763	144763	144763	144763	144763	144763	144763	144763	144763	144763	144763	144763
Number of countries	26	26	26	26	26	26	26	26	26	26	26	26
Country fixed effects	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No

Note: * and ** denote statistical significance at the 10% and 5% levels, based on heterogeneity robust standard errors. Individual-level variables used in the regressions are aggregated at the school level.